



Colorado Assessment Implementation Study

Prepared for
The Colorado Department of Education

May 2014

Disclaimer

This study, focused on the implementation of assessments in Colorado, is a direct response to concerns raised by the field. It is neither a study about the Colorado Academic Standards and consortia-developed assessments nor a cost-benefit analysis. This work was originally produced in whole or in part by WestEd with funds from the U.S. Department of Education under cooperative agreement number S283B120016. The content does not necessarily reflect the position or policy of the Department of Education, nor does mention or visual representation of trade names, commercial products, or organizations imply endorsement by the federal government.

About WestEd

WestEd is a non-partisan, non-profit, research, development, and service agency that works with education and other communities to promote excellence, achieve equity, and improve learning for children, youth, and adults. WestEd provides consulting and technical assistance, evaluation, professional development, and research and development services to schools, districts, state education agencies, and education organizations across the country.

Table of Contents

EXECUTIVE SUMMARY	i
I. INTRODUCTION	1
Purpose of Study	1
Federal Statute	1
State Statute	2
Assessment Overview	2
State Demographic Data	2
II. METHODOLOGY	3
Design	3
Colorado Assessment Implementation Survey	3
Focus Groups	3
Research Constraints and Limitations	4
III. RESULTS	5
Overall Findings	5
Survey Findings	5
Focus Groups Findings	14
IV. DISCUSSION	18
Implementation Approaches	18
Future Directions	21
V. APPENDICES	23
A. Colorado Assessment Implementation Survey	23
B. Focus Group Protocol and Questions	30
C. Focus Group Summary Tables	33
D. Feedback from the Commissioner's Superintendent Advisory Council	48

LIST OF TABLES

Table 1. Survey Response Rate in Rural, Suburban, and Urban Areas by Region and Student Population	6
Table 2. Perceptions of Factors Affecting Readiness to Administer the State Assessments	8
Table 3. Perceptions of the Value and Burden of State and Local Assessments	9
Table 4. Assessments Ranked by Perceived Value	10
Table 5. Assessments Ranked by Perceived Burden	10
Table 6. Important Characteristics of a State Assessment System	12

LIST OF FIGURES

Figure A. Participation in Field Tests and Educator Effectiveness Pilot	7
Figure B. Burden versus Value of Assessments for Informing Student Progress	11
Figure C. Burden versus Value of Assessments for Improving School/District Performance	11
Figure D. Focus Group Participants	14

EXECUTIVE SUMMARY

Introduction

The Colorado State Assessment System is designed to measure student mastery of the state academic content standards. Revised standards were adopted in December 2009 and August 2010. Since that time several new tests have been introduced. In 2012 the state implemented the Transitional Colorado Assessment Program (TCAP) to reflect changes in the academic content standards and provide better information to educators. As the state assessment system is refined and continues to evolve, intended and unintended outcomes may arise for educators, students, and parents. These outcomes become complicated when viewed in light of early literacy assessments, local district assessment systems, new online assessments, and additional education reform initiatives. To address concerns raised by educators, parents, and other stakeholders, the Colorado Department of Education (CDE) decided to gather information about the implications of the entire assessment system. Lacking the resources for such a study, CDE turned to its Regional Comprehensive Center for assistance.

CDE worked with the Regional Comprehensive Center to design a multi-phase study at no cost to CDE or the state. The purpose of the study is to discern and examine issues and concerns associated with implementation of the new state assessment system, and provide feedback that informs policy, practice, and future directions. The first phase of the study captured perceptions and sentiments about current and new state tests, challenges and needs associated with transitioning to the new assessment system, proposed solutions that address key challenges, and ideas for implementing a high-quality assessment system.

Methodology

Conducted in two phases, this report includes findings from phase one, which took place between February and April 2014. The first phase involved a review of documents and district artifacts, survey of district assessment coordinators, and district and role-alike focus groups. The second phase of the study began in May and includes a brief, follow-up survey of district assessment coordinators; follow-up conversations with district focus groups; and a focus group of large, metro-area districts.

Survey

A voluntary survey was sent to 178 district assessment coordinators (DAC) on March 12. DACs were encouraged to consult other district leaders to ensure that responses reflected those of the district as a whole. The survey had three sections. The first section asked for demographic information to assist in the analyses. The second section allowed respondents to provide feedback on general readiness issues related to the new assessment system. The third section requested information on the value and burden of state and other assessments. Each district could submit one completed survey. The survey closed March 28.

Focus Groups

CDE sent invitations to every district during the last week of January 2014. Twenty-three districts responded to the request. The Commissioner's Superintendent Advisory Council helped to select the following districts from the list of 23 respondents: Archuleta, Buena Vista R-31, Cherry Creek 5, Delta County, La Veta RE-2, Platte Valley RE-7, Strasburg, and Woodland Park Re-2. CDE also identified teachers from across the state to participate in a separate focus group and reached out to various concerned parent groups as well as the Colorado League of Charter Schools to identify participants for two additional role-alike focus groups.

Research Constraints and Limitations

The state assessment system is the primary focus of the study. Study participants also raised accountability issues, such as whether the results of new tests are sufficiently valid and reliable to support school accountability and teacher evaluation decisions. These issues were noted and are discussed in the report. For analyses of the survey data, districts were weighted equally. Since all districts, regardless of location or size, need to perform similar transition activities, each had an equal voice during survey administration and analysis. As a consequence, however, it should be noted that the views of rural districts with small student populations have a proportionally higher impact on the results.

Results

Eighty-seven DACs submitted complete surveys for a 49% response rate. Surveys were submitted by districts in each region with the highest rates for the northeast (23%) and southwest (20%) regions. Districts were predominantly rural (73%) with suburban and urban districts accounting for 16% and 8% respectively of the sample. More than half of the districts are small with 54% serving less than 1,000 students. A majority (78%) of districts participated in assessment field tests (i.e., new English language arts and/or mathematics tests, science and/or social studies tests) and/or the educator effectiveness pilot. Ninety-three individuals participated in focus groups with a majority (nearly 80%) representing rural districts. Twenty percent of focus group participants were parents, and another 20% were teachers. Eleven percent of participants were principals, 11% were either superintendents or assistant superintendents, and an additional 11% represented charter schools.

Cross-Cutting Themes

Analyses revealed the following areas of consensus across survey respondents and focus group participants regarding the state assessment system:

Value Most — Participants value local interim assessments more than statewide summative assessments. State assessment results and the information provided in reports, however, as well as the emphasis on growth are valued. Additionally, participants value the common metric and basis the assessments provide for the state's accountability system.

Value Least — Participants dislike the high stakes associated with state assessment results. They report that results are delayed and general, and do not inform instructional practice, programming, and student placement decisions (i.e., utility). They also question the timing of statewide summative tests that are administered February through early April rather than at the end of the school year in May.

Challenges — Time is the theme that summarizes the challenges associated with implementation of the new assessment system. Issues related to this theme include the impact on instructional time, frequent schedule disruptions, computer and device access for non-testing students during test windows, and the burden associated with test preparation and administration.

Needs — To facilitate the transition to the new assessment system, participants identified the need for efficient methods to administer tests while minimizing disruptions to instructional time, sufficient numbers and types of devices to meet instructional needs and technical requirements, and enhanced staffing and capacity to support the process.

Solutions — Proposed solutions cluster under the themes of flexibility and results. Flexibility involves options based upon district needs, size, and location as well as local decisionmaking regarding how often as well as whether, when, where, and how students are tested. Study participants also expect timely, high-quality, relevant, and useful feedback about student performance that informs educators, parents, and students.

Unique Survey Results

Respondents (80%) value local interim assessments most followed by early literacy assessments associated with the READ Act (50%). More than half (57%) of urban districts consider their local interim assessments as high burden compared to 37% of rural and 23% of suburban districts. Few respondents (21%) value the social studies and science assessments. School readiness assessments impose the greatest burden (76%) followed by social studies (74%) and science (73%) tests. The least burdensome tests include the Colorado ACT and other district postsecondary readiness assessments. These tests, however, are identified as high burden to half of the urban districts compared to 19% of rural and none of the suburban districts.

Districts generally view current English language arts and mathematics assessments (i.e., TCAP) as low in value with suburban districts valuing these assessments more than urban and rural districts. Two-thirds of rural and suburban districts view all TCAP assessments as high in burden compared to 29% of urban districts. Urban districts (50%) value the science assessments at higher levels than suburban (38% believe the tests inform student progress, 25% say they support school improvement) and rural districts (20%). And although the burden associated with school readiness assessments is high, urban districts (100%) consider it highly valuable in informing student progress compared to 38% of rural and 13% of suburban districts.

Regarding overall readiness to administer state assessments, 28% of districts appear fully prepared and 20% are not yet ready. The two primary issues that influence district readiness are management (62%) and devices (60%). This holds for rural districts with 63% citing management and 57% noting devices. Devices have a greater impact on suburban (71%) and urban (71%) readiness. Suburban districts appear to be the least prepared with 79% citing management and the capacity of their information technology staff. Also, nearly two-thirds of suburban districts reported network infrastructure challenges compared to 49% of rural and 43% of urban districts.

Unique Focus Group Results

In addition to the cross-cutting themes, each role group articulated unique concerns and potential solutions. Students, for example, fear that new exams will be challenging, include unfamiliar content, and be unfair or inadequate measures for some students. They worry about fatigue given the amount of time spent viewing a computer screen during test sessions, and want simpler tests with

one section per subject area. Parents dislike the pressure placed on students, want more transparency about test items and the assessment process, and prefer fewer summative tests. Teachers are familiar with the current system and expressed fears about moving to a new and different system. They also dislike the stress that high-stakes tests place upon students and teachers. Principals value the READ Act because they may select tests from an approved list, and useful, diagnostic information is obtained quickly. Assessment coordinators fear that new tests will not be user-friendly and view social studies exams as problematic due to timing and lack of incentives for high school seniors. Technology directors hope that their districts are well prepared to administer new online assessments and view feasibility (i.e., capacity, complexity, scale, resources, and timeframe) as a major challenge. And superintendents raise accountability issues and concerns about how assessment results will be used. They question the quality of new tests, want fair and accurate measures that reflect student learning, and view communication about the process and results as a challenge.

A solution noted by focus group participants is that of holding schools and districts harmless until all components of the system are functioning effectively and validated. Flexibility is another key theme. Participants, especially those from rural districts, want more local control and differentiated options based upon district needs, size, and location. Funding flexibility to repair and improve school facilities, support teachers, upgrade technology, and meet other deferred needs was also mentioned. Finally, because participants feel overwhelmed and under-resourced, they desire a more gradual pace and seek to slow the roll-out of the new assessment system.

Conclusion

Phase one of this study examined the issues associated with implementation of the state assessment system. While differences were observed among districts located in rural, suburban, and urban areas, several cross-cutting themes and common challenges emerged:

- ❖ Impact on instructional time
- ❖ Moderate levels of readiness (i.e., management, devices, and capacity)
- ❖ Quantity, frequency, and length of assessments
- ❖ Need for timely, relevant, and useful results
- ❖ Burden and utility of assessments at the elementary and secondary levels
- ❖ Recognition of local assessment systems and practices

Based upon the findings four implementation approaches are proposed for consideration:

1. Stay the course and implement the transition plan as scheduled
2. Stay the course with added supports and policy adjustments
3. Purposefully delay parts of the system
4. Strategically eliminate specific assessments

Study findings and potential solutions that address common challenges were discussed among members of the Commissioner's Superintendent Advisory Council on May 1, 2014. The council's feedback informed the implementation approaches and helped clarify options for minimizing the assessment burden on districts. Several council members expressed an interest in limiting statewide summative assessments to the federal minimum and making optional any assessments beyond that minimum.

The second phase of the study began in May following online administration of the new science and social studies assessments and field testing of the new online English language arts and mathematics assessments. The objective of the second phase is to understand whether and how the challenges and opportunities may have changed, gather lessons from the state's first online administration, and solicit feedback on strategies for facilitating a smooth transition to the new assessment system. The Colorado Department of Education intends to use the findings, input, and additional feedback to address unintended consequences of the new assessment system and inform adjustments to administrative policies and procedures.

I. INTRODUCTION

The Colorado State Assessment System is designed to measure student mastery of the state academic content standards. Revised standards were adopted in December 2009 and August 2010. Since that time several new tests have been introduced. In 2012 the state implemented the Transitional Colorado Assessment Program (TCAP) to reflect changes in the academic content standards and provide better information to teachers as they began aligning curricula and instruction with the revised standards. That same year Colorado became a governing member of the Partnership for Assessment of Readiness for College and Careers (PARCC), which is developing assessments in English language arts and mathematics for grades 3–8 and high school. The PARCC assessments are scheduled for implementation in 2015. Other new tests added to the system include the ACCESS for ELLs® English Language Proficiency assessment implemented during the 2012–2013 school year and general and alternate assessments for science and social studies implemented this spring.

Purpose of Study

As the state assessment system is refined and continues to evolve, some intended and unintended outcomes may arise for educators, students, and parents. These outcomes become more complex when viewed in light of early literacy assessments, local district assessment systems, new online assessments, and other education reform initiatives. To address concerns raised by educators, parents, and other stakeholders, the Colorado Department of Education (CDE) decided to gather information about the implications of the entire assessment system. Lacking the resources for such a study, CDE turned to its Regional Comprehensive Center for assistance. CDE worked with the Regional Comprehensive Center to design a multi-phase study at no cost to CDE or the state. The purpose of the study is to discern and examine issues and concerns associated with implementation of the new state assessment system, and provide feedback to CDE that informs policy, practice, and future directions.

Federal Statute

States that accept federal funds for purposes such as supporting the education of children living in poverty, English language learners, and students with disabilities are required to administer statewide assessments to all students. Currently, Colorado receives approximately \$326 million in federal funds for these and related purposes and is therefore required to administer the following assessments:

- ❖ English language arts and mathematics for grades 3 through 8;
- ❖ English language arts and mathematics at least once in high school; and
- ❖ Science at least once in elementary, middle, and high school.

For the aforementioned assessments, states must give the same assessments to all students and at least 95% of the students must participate. There are also some required assessments specific to certain populations of students such as English language learners (ELL).

State Statute

As outlined in the state's revised statutes, Colorado's new statewide assessment system includes the following tests:

- ❖ English language arts for grades 3 through 11;
- ❖ Math for grades 3 through 8 and three times in high school;
- ❖ Science at least once in elementary, middle, and high school;
- ❖ Social studies once in elementary, middle, and high school;
- ❖ ACT in grade 11;
- ❖ WIDA ACCESS for ELLs; and
- ❖ State-mandated, locally-determined assessments for the READ Act and school readiness.

Assessment Overview

Colorado assessments are changing in order to accurately assess student mastery of the new Colorado Academic Standards. As standards become more coherent and rigorous, assessments must adapt to align with them. As a result, the Transitional Colorado Assessment Program is being phased out and replaced by the Colorado Measures of Academic Success (CMAS): the state's new English language arts, mathematics, science, and social studies assessments.

Colorado is a governing member of a multi-state assessment consortium called the Partnership for Assessment of Readiness for College and Careers (PARCC). Beginning in 2014–2015, the online PARCC assessments will be administered in grades 3 through 11 for English language arts and grades 3 through 8 and three years of high school for mathematics. Additionally, for the first time this school year, new online state science and social studies assessments were administered. Science assessments are given in grades 5, 8, and 12, while the social studies assessments are administered in grades 4, 7, and 12. The grade 12 high school assessments are scheduled for the fall of 2014.

State Demographic Data

The state of Colorado has 178 school districts, excluding Boards of Cooperative Educational Services (BOCES), charter schools, and detention centers. Districts are located throughout eight regions: Metro, Pikes Peak, North Central, Northeast, Northwest, Southeast, Southwest, and West Central, with the greatest concentration in the Northeast region ($n=32$). Overall, no more than 18% of districts are located in any one region. While the Metro region contains only 18 districts, making it the second smallest region in terms of the number of districts, this region serves more than half (55%) of the student population in the state. Conversely, the Northeast has the greatest number of districts yet serves less than 2% of the state's student population. Additionally, more than half (60%) of the districts serve 1,000 students or fewer, 11% serve more than 20,000 students, and the 15 largest districts account for 68% of the state's student population.

II. METHODOLOGY

The study is being conducted in phases to capture information, lessons learned, and options for improvement throughout implementation of the new assessment system. Researchers utilized a statewide district-level online survey and targeted focus groups (at least one per region). Both the focus groups and survey included questions about the components of local and state assessment systems and the impact of current and new assessments.

Design

The first phase of the study took place between February and April 2014. Three approaches were used to gather information: (1) document and artifact review, (2) survey of district assessment coordinators, and (3) focus groups. Phase two began in May, continues until July, and involves similar approaches: (1) follow-up survey of district assessment coordinators; (2) follow-up conversations with the district focus groups; (3) focus group of large, metro-area districts; and (4) a phone interview with online education providers.

Colorado Assessment Implementation Survey

A voluntary survey (See Appendix A) — created by WestEd researchers in collaboration with CDE staff — was sent to 178 district assessment coordinators (DAC) on March 12, 2014. DACs were encouraged to consult other district leaders to ensure that responses reflected those of the district as a whole. The purpose of the survey was to gather perceptions regarding the value and burden of state-required and other assessments (e.g., locally administered interim assessments) as well as views of general readiness issues related to the new assessment system.

Survey responses were reviewed and incomplete surveys eliminated from the sample. Rating questions were reclassified as either low (i.e., low, somewhat low, moderate) or high (i.e., somewhat high and high) to achieve sufficient counts for chi-square analyses. This allowed researchers to better understand how districts in different areas (i.e., rural, suburban, and urban) perceive the value and burden of current and new assessments. Additional chi-square analyses were conducted using student population as a frame of reference. Districts were reclassified by student population as follows:

- ❖ Small = 1–250 and 251–1,000 students
- ❖ Medium = 1,001–3,000 and 3,001–10,000 students
- ❖ Large = 10,001–20,000 and more than 20,000 students

Statistically significant differences were not found across responses for small, medium, and large districts, and thus are not reported in the results section.

Focus Groups

To delve deeper into assessment implementation challenges and opportunities, WestEd researchers convened eleven focus groups. CDE sent invitations to every district during the last week of January 2014. Twenty-three districts responded to the request. The Commissioner's Superintendent Advisory Council helped to select the following districts from the list of 23 respondents: Archuleta, Buena Vista R-31, Cherry Creek 5, Delta County, La Veta RE-2, Platte Valley RE-7, Strasburg, and Woodland

Colorado Assessment Implementation Study

Park Re-2. Each district focus group consisted of approximately 7–10 participants, including superintendents, assessment coordinators, curriculum coordinators, technology directors, principals, teachers, parents, and students, all of whom were selected by the districts. CDE also identified teachers from across the state to participate in a separate focus group and reached out to concerned parent groups (e.g., Gifted Education Advisory Council, Padres Unidos, Speak for Cherry Creek, Special Education Advisory Council, Stand for Children, State Council for Educator Effectiveness, and groups opposed to the Common Core State Standards) as well as the Colorado League of Charter Schools to identify participants for two additional role-alike focus groups. (See Appendix B)

Conducted between February 26 and April 2, each focus group convened for approximately two hours. Researchers posed 10 questions organized as follows: current system (value most and least); new system (hopes, fears, challenges); and managing the transition (needs and solutions). Comments and themes were summarized on flip charts or whiteboards during each session. Notes were also typed during the sessions and commenter roles indicated. For each district and role group, participant comments were reviewed, organized into eight categories, coded by theme, and tallied. (See Appendix C)

Research Constraints and Limitations

The state assessment system is the primary focus of the study. Study participants also raised accountability issues, such as whether the results of new tests will be sufficiently valid and reliable to support decisions regarding school accountability and teacher evaluation. These issues were noted and are described in the discussion section of the report. For analyses of the survey data, districts were weighted equally. Since all districts, regardless of location or size, need to perform similar transition activities, each had an equal voice during survey administration and analysis. As a consequence, however, it should be noted that the views of rural districts with small student populations have a proportionally higher impact on the results.

III. RESULTS

Findings from the first phase of the assessment implementation study follow. Information was collected from documents and artifacts, a statewide survey of district assessment coordinators, and focus groups of a representative sample of districts and various stakeholders.

Overall Findings

Analyses revealed the following areas of consensus across survey respondents and focus group participants regarding the state assessment system:

Value Most — Participants value local interim assessments most. They also appreciate the format, elements, and feedback contained in state assessment reports as well as the emphasis on growth. Additionally, the common metric and basis the assessments provide for the state’s accountability system are valued.

Value Least — Participants dislike the high stakes associated with state assessment results. They report that results are delayed and general, and do not inform instructional practice, programming, and student placement decisions (i.e., utility). They also question the timing of statewide summative tests that are administered February through early April rather than at the end of the school year in May.

Challenges — Time is the theme that summarizes the challenges associated with implementation of the new assessment system. Issues related to this theme include the impact on instructional time, expanded testing calendar, frequent schedule disruptions, computer and device access for non-testing students during testing windows, and the burden associated with test preparation and administration.

Needs — To facilitate the transition to the new assessment system, participants identified the need for efficient methods to administer tests while minimizing disruptions to instructional time, sufficient numbers and types of devices to meet instructional needs and technical requirements, and enhanced staffing and capacity to support the process.

Solutions — Proposed solutions clustered under the themes of flexibility and results. Flexibility involves options based upon district needs, size, and location as well as local decisionmaking regarding how often as well as whether, when, where, and how students are tested. Study participants also expect timely, high-quality, relevant, and useful feedback about student performance that informs educators, parents, and students.

Survey Findings

Demographics

Eighty-seven DACs submitted complete surveys for a 49% response rate. Surveys were submitted by districts in each region. Table 1 presents district participation in rural, suburban, and urban areas by region and student population. Overall, about three quarters of the respondents are from rural areas. The Northeast region is well represented in the survey sample. Specifically, 24% of participating districts are from that region. Additionally, districts serving one to 1,000 students are highly represented (54%) in the survey sample. On an area by area basis, at least half of the suburban and urban districts are located in the Metro region. Rural districts are located in every region except for the Metro area.

Colorado Assessment Implementation Study

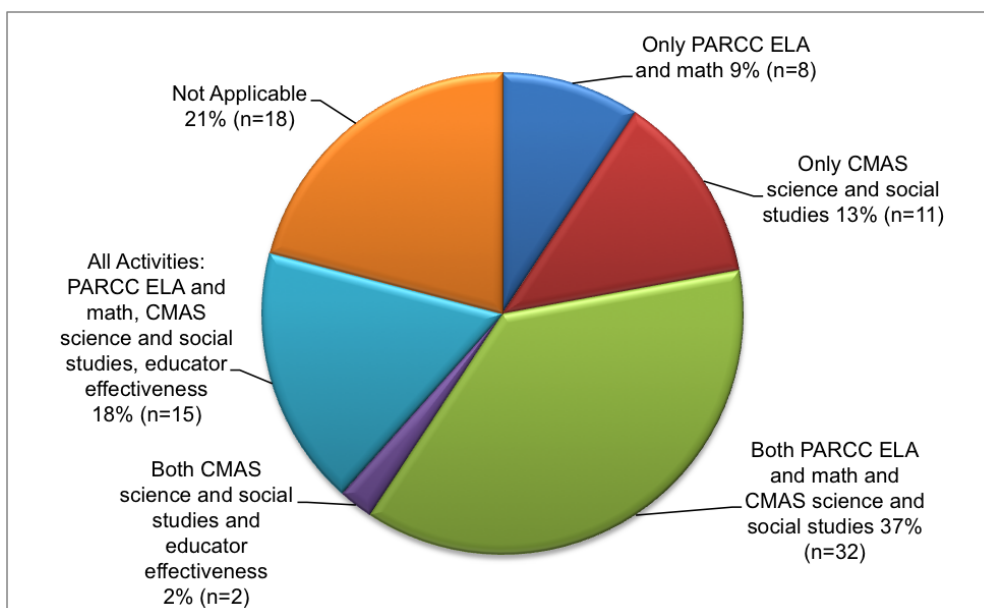
More than two-thirds of rural districts are relatively small, serving between one and 1,000 students. About 71% of suburban and 57% of urban districts serve more than 10,000 students compared to 2% of rural districts.

Table 1. Survey Response Rate in Rural, Suburban, and Urban Areas by Region and Student Population (n=87)

	Rural (%)	Suburban (%)	Urban (%)	Overall (%)
Overall	76	16	8	100
Region				
Metro	0	50	57	13
Pikes Peak	2	29	29	8
North Central	6	21	0	8
Northeast	31	0	0	24
Northwest	5	0	14	5
Southeast	20	0	0	15
Southwest	26	0	0	20
West Central	11	0	0	8
Student Population				
1–1,000	68	7	14	54
1,001–10,000	31	21	29	29
More than 10,000	2	71	57	17

A majority (78%, n=68) of districts participated in assessment field tests (i.e., new English language arts and/or mathematics tests, science and/or social studies tests) and/or the educator effectiveness pilot. Just 18 districts indicated that they did not participate in any of these activities. More than one third of participating districts reported that they will or have participated in both PARCC English language arts/mathematics and CMAS science/social studies field tests. Additionally, 18% of districts participated in all of the activities: PARCC and CMAS field tests as well as the educator effectiveness pilot.

Figure A. Participation in Field Tests and Educator Effectiveness Pilot



Readiness

Table 2 presents districts' perceptions of factors affecting their readiness to administer online state assessments. The percentages for each area (i.e., rural, suburban, and urban) and overall indicate that the factor has a somewhat high or high impact on their readiness. Regarding overall readiness, 28% of districts appear fully prepared. And although not displayed in Table 2, about half of the districts are moderately prepared while 20% are not yet ready. Two issues affect district readiness the most: management (e.g., planning, organizing, scheduling, logistics, administering, controlling) and devices (e.g., hardware and software at schools, device-to-student ratios), which influence 62% and 60% of districts respectively. This holds for rural districts with 63% citing management and 57% noting devices. Devices have a greater impact on suburban (71%) and urban (71%) readiness.

Suburban districts report that the factors most impacting their readiness are management and information technology staff with 79% citing these issues. Also, nearly two-thirds of suburban districts reported network infrastructure concerns compared to 49% of rural and 43% of urban districts. Interestingly, devices have less of an influence on readiness for rural districts (57%) than for suburban (71%) and urban (71%) districts. Management is less of an influence, however, on urban districts' readiness compared to suburban and rural districts. Only 29% of urban districts indicated that their readiness is highly impacted by management compared to 63% of rural and 79% of suburban districts. In addition, few (17%) urban districts reported that training and professional development are issues impacting their readiness compared to half of suburban and rural districts.

Table 2. Perceptions of Factors Affecting Readiness to Administer the State Assessments (n=87)

	Rural (%)	Suburban (%)	Urban (%)	Overall (%)
Overall readiness	30	14	29	28
Instructional design and preparation (e.g., curricular materials and resources, instructional strategies)	42	57	57	45
Network infrastructure (e.g., Internet access, wired and wireless connectivity)	49	64	43	51
Devices (i.e., hardware and software at schools, device-to-student ratios)	57	71	71	60
Management (e.g., planning, organizing, scheduling, administering, controlling)	63	79	29	62
IT staff and personnel	47	79	57	52
Procurement of necessary funding & resources	44	50	57	46
Training/professional development	48	50	17	46
Notes: The percentages above indicate that the factor has a somewhat high or high impact on their readiness. Bold figures represent large, though not necessarily statistically significant, differences across sectors.				

Value and Burden

Table 3 presents participating districts' perceptions of the value and burden of state and local assessments. Respondents (80%) value local interim assessments most followed by early literacy assessments associated with the READ Act (50%). More than half (57%) of urban districts consider their local interim assessments as high burden compared to 37% of rural and 23% of suburban districts. School readiness assessments impose the greatest burden (76%) followed by social studies (74%) and science (73%) tests. The least burdensome tests include the Colorado ACT and other district postsecondary readiness assessments. These tests are identified, however, as high burden to half of the urban districts compared to 19% of rural and none of the suburban districts. District rankings of the value (highest to lowest) and burden (least to most) of assessments are provided in Tables 4 and 5. These tables are shaded green, yellow, orange, and red. Green indicates the tests that are most valued and of least burden, and red indicates the tests that are least valued and of greatest burden. Yellow and orange shading indicates tests that fall between these distinctions.

Districts generally view current English language arts and mathematics assessments (i.e., TCAP) as low value with suburban districts valuing these assessments more than urban and rural districts. Two-thirds of rural and suburban districts view all TCAP assessments as high in burden compared to 29% of urban districts. Few respondents (21%) value the social studies and science assessments. Urban districts (50%) value the science assessments at higher levels than suburban (38% believe the tests inform student progress, 25% say they support school improvement) and rural districts (20%). And although the burden of school readiness assessments is high, urban districts (100%) consider it highly valuable in informing student progress compared to 38% of rural and 13% of suburban districts. The urban sample size for this item, however, only reflects two districts.

Table 3. Perceptions of the Value and Burden of State and Local Assessments

	Value				Assessment (State or Local)	Burden			
	Rural (%)	Suburban (%)	Urban (%)	Overall (%)		Rural (%)	Suburban (%)	Urban (%)	Overall (%)
Informs Student Progress	27	43	29	29	TCAP Reading (n=85)	70	79	30	67
Improves School/District	33	36	29	33					
Informs Student Progress	23	43	29	27	TCAP Writing (n=85)	70	79	29	67
Improves School/District	28	29	29	28					
Informs Student Progress	24	43	29	27	TCAP Mathematics (n=83)	66	79	29	64
Improves School/District	32	39	29	33					
Informs Student Progress	20	38	50	23	CMAS Science (n=69)	72	82	80	73
Improves School/District	20	25	50	21					
Informs Student Progress	18	38	33	21	CMAS Social Studies (n=69)	71	82	100	74
Improves School/District	20	25	33	21					
Informs Student Progress	31	57	57	39	ACCESS for ELLs (n=70)	61	57	57	59
Improves School/District	27	39	29	29					
Informs Student Progress	44	43	50	44	Colorado ACT (n=81)	18	14	17	17
Improves School/District	43	43	33	42					
Informs Student Progress	64	64	57	64	Early Literacy Assessments READ Act (n=80)	64	71	57	65
Improves School/District	54	36	43	51					
Informs Student Progress	38	13	100	36	School Readiness Assessment (n=46)	72	88	100	76
Improves School/District	38	13	0	34					
Informs Student Progress	90	77	86	87	Other District Interim Assessments (n=79)	37	23	57	36
Improves School/District	83	62	86	80					
Informs Student Progress	61	29	50	55	Other District Postsecondary Readiness (n=41)	19	0	50	17
Improves School/District	47	29	50	44					

Table 4. Assessments Ranked by Perceived Value

Rank	Assessment	Informs Student Progress	Improves School/District
1	District administered interim	87 %	80 %
2	Early Literacy (READ Act)	64 %	51 %
3	District administered postsecondary readiness	55 %	44 %
4	Colorado ACT	44 %	42 %
5	ACCESS for ELLs	39 %	29 %
6	School Readiness	36 %	34 %
7	TCAP Reading	29 %	33 %
8	TCAP Mathematics	27 %	33 %
9	TCAP Writing	27 %	28 %
10	CMAS Science	23 %	21 %
11	CMAS Social Studies	21 %	21 %

Table 5. Assessments Ranked by Perceived Burden

Rank	Assessment	Burden
1	District administered postsecondary readiness	17 %
1	Colorado ACT	17 %
3	District administered interim	36 %
4	ACCESS for ELLs	59 %
5	TCAP Mathematics	64 %
6	Early Literacy (READ Act)	65 %
7	TCAP Reading	67 %
7	TCAP Writing	67 %
9	CMAS Science	73 %
10	CMAS Social Studies	74 %
11	School Readiness	76 %

Figures B and C are scatter plots comparing the burden (x-axis) to the value (y-axis) of assessments for informing student progress and improving school/district performance. Tests appearing in the upper right-hand quadrant are viewed as high burden and high value. Tests appearing in the upper left-hand quadrant are viewed as low burden and high value. Most of the tests are clustered in the lower right-hand quadrant indicating high burden and low value.

Figure B. Burden versus Value of Assessments for Informing Student Progress

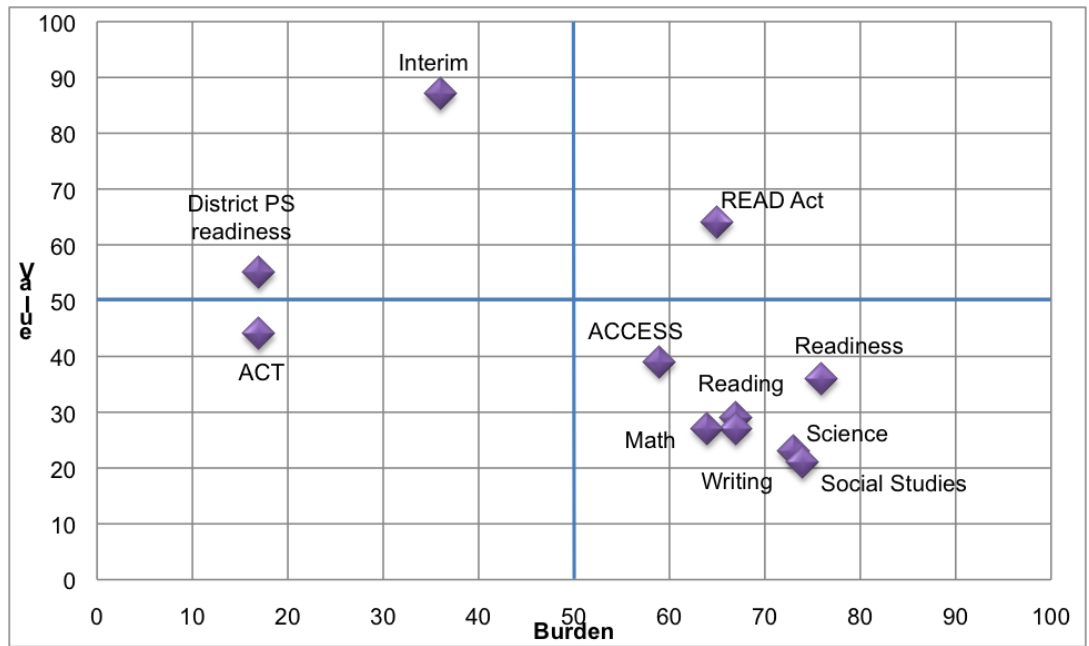
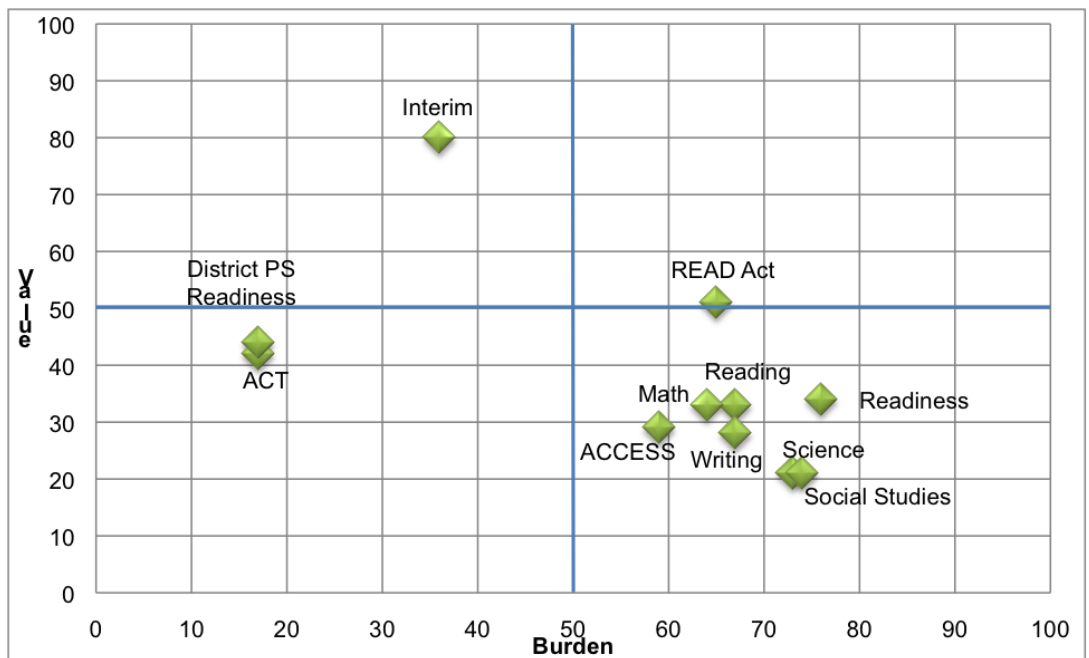


Figure C. Burden versus Value of Assessments for Improving School/District Performance



Elements of a Statewide System

The survey asked districts to rate the importance of particular elements in a state assessment system. As indicated in Table 6, the most important characteristic for rural, suburban, and urban districts alike is timeliness of assessment results (93%). The second and third most desired characteristics are information on student growth from year to year (80%) and length of assessments (74%). On the other hand, only 20% of participating districts reported that assessing social studies annually in grades 4, 7, and 12 is an important element.

Perceptions vary by area and reveal notable differences. For example, cross-school comparisons are highly important to 71% of suburban districts compared to 43% of urban and 21% of rural districts ($p < .001$). Additionally, cross-district comparisons are more important to urban (57%) and suburban (64%) districts compared to 14% of rural districts ($p < .001$). Indicators of school readiness are far more important to urban districts (86%) compared to 23% of suburban and 41% of rural districts ($p < .05$).

Table 6. Important Characteristics of a State Assessment System

	Rural (%)	Suburban (%)	Urban (%)	Overall (%)
Timeliness of results	90	100	100	93
Information on student growth from year to year	77	86	86	80
Length of assessments	73	69	86	74
Actionable information at the program level	62	79	86	68
Actionable information at the student level	61	79	86	67
Aligned local and state assessment system	67	64	71	66
Flexible state assessment window	64	64	71	65
Inclusion of writing	60	64	71	61
Information on student mastery of the Colorado Academic Standards	53	79	71	60
Items beyond selected response	54	64	100	59
Early indicators of college and career readiness	50	64	71	55
Indicators of early literacy development	53	57	57	54
Indicators of school readiness*	41	23	86	43

	Rural (%)	Suburban (%)	Urban (%)	Overall (%)
Gradual transition from paper to online administration	42	29	43	40
Single state assessment window	35	31	14	32
Cross-school comparisons**	21	71	43	31
Cross-state comparisons	22	36	57	27
Cross-district comparisons**	14	64	57	26
Assessing social studies annually in grades 4, 7, and 12	24	7	14	20
*p<.05; **p<.001				

Respondents also provided feedback on the elements valued most and least about the current system, concerns regarding the new system, and the types of assistance that would increase readiness and facilitate the transition from current to new.

The top three items, selected from a list of 15, that participating districts value most about the current assessment system are that it provides:

- ❖ A common basis for the state's accountability system (n=29);
- ❖ Feedback to educators (n=28); and
- ❖ Comparisons across schools and districts (n=28).

On the other hand, the top three items that participating districts value least about the current assessment system are the following:

- ❖ Inform instructional practice (n=23);
- ❖ Identify teacher development needs (n=22); and
- ❖ Ensure accountability for student performance and progress (n=18).

As the state moves forward with the new assessment system, districts' prominent concerns are:

- ❖ Impact on instructional time (n=63);
- ❖ Quantity of assessments (n=46);
- ❖ Time required for administration (n=40); and
- ❖ Technology requirements (n=35).

To facilitate the transition from the current to the new system, districts need the following types of assistance:

- ❖ Efficient methods to administer assessments that minimize disruptions to instruction time (n=53);
- ❖ Procuring devices that meet instructional needs and assessment consortia requirements (n=46); and
- ❖ Professional development for staff to understand and use assessment results (n=28).

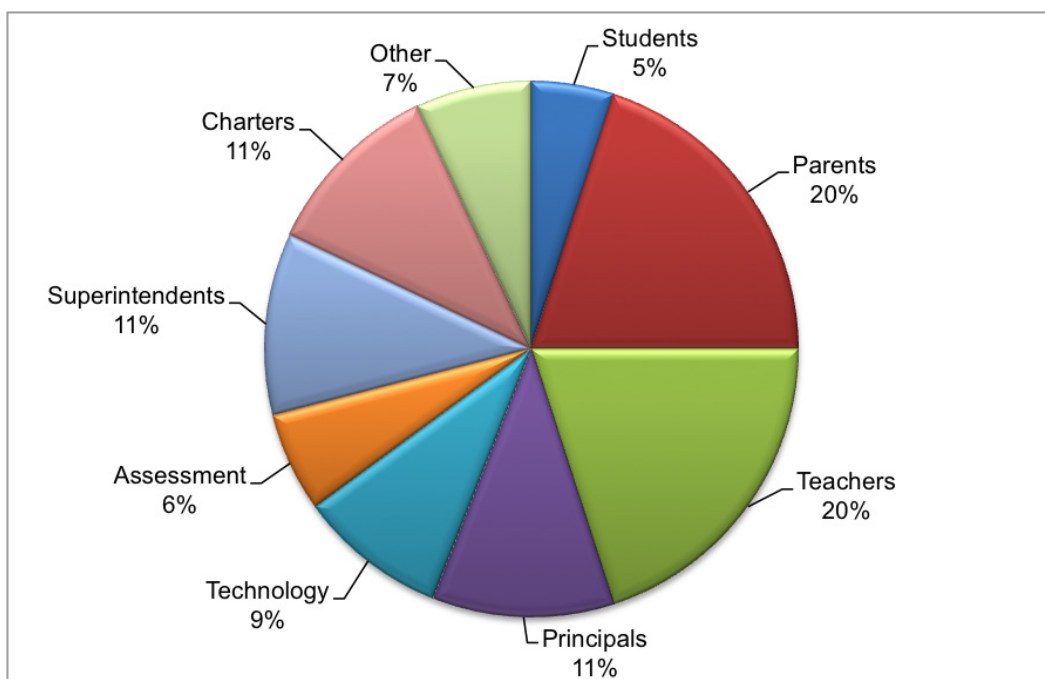
Focus Groups Findings

The focus group sessions provided superintendents, assessment coordinators, technology directors, principals, teachers, parents, students, and other stakeholders with an opportunity to share challenges, needs, and proposed solutions for implementing a high-quality assessment system. They also described how an ideal state assessment system might look in the future.

Demographics

Ninety-three individuals participated in focus groups with a majority (nearly 80%) representing rural districts. Twenty percent of focus group participants were parents, and another 20% were teachers. Eleven percent of participants were principals, 11% were either superintendents or assistant superintendents, and another 11% represented charter schools. Participating role groups are depicted in Figure D.

Figure D. Focus Group Participants



Current State Assessment System

Value Most — Participants value the Colorado ACT, especially students, parents, and secondary educators. ACT results are timely, relevant, and meaningful to students and their families. The current TCAP assessments provide a common metric that enables comparisons of schools and districts across the state. Participants also value the reports generated from assessment results and emphasis on student growth. They like the data and graphs depicting students' performance levels, weaknesses, and strengths. And they appreciate the user-friendly format, which allows students, parents, and teachers to view trends over time.

Value Least — Test results are delayed and thus do not inform scheduling, student placements, or instructional decisions. Test administration is perceived as a burden because it diminishes the

amount of time available for instruction and other educational purposes. Participants dislike the weight and number of accountability indicators associated with TCAP assessment results. While the stakes are high for schools and teachers, there are no consequences for students. Few students demonstrate support and buy-in for state summative assessments, and feel limited responsibility to perform well on the assessments.

New State Assessment System

Hopes — Participants hope that new online assessments will be user-friendly, interactive, and relevant, so that students will be engaged and motivated to perform well. They expect results in a timely manner so that educators may use the information to immediately inform instruction and determine necessary interventions. Online administration should reduce the overall amount of time spent testing as well as the costs associated with preparation and administration.

Fears — Students, particularly those at the elementary level, may not be adequately prepared for online assessments. Participants believe that young students lack keyboarding and other technological skills deemed necessary for successful completion of online assessments. This highlights a common fear that the new assessments will measure students' technology skills rather than content knowledge and application.

Challenges — New assessments will be administered more frequently and further reduce instructional time, especially for students who need it the most (e.g., English language learners and students with disabilities). Furthermore, expanded testing windows and increased testing time will limit access to computer labs and devices for non-testing students and teachers who normally use the facilities for their classes. Districts may not have the capacity to efficiently transition to the new system. Devices, network infrastructure (e.g., reliable bandwidth), and other resources are insufficient in some areas. Several sites lack the staff capacity to simultaneously administer assessments and provide instruction to students who are not testing.

Needs — Participants desire comprehensive professional development that prepares staff to administer the assessments effectively, and understand and apply the results. Additional professional development opportunities are essential to help educators better understand and align formative assessments, instructional strategies, and curricula to the Colorado Academic Standards. Participants also indicated the need for sufficient numbers of compatible devices as well as adequate and well maintained network infrastructures. Educators requested resource banks that contain curricular materials, sample items, and model units, to help prepare students for the new online assessments.

Solutions — A solution noted frequently by focus group participants is that of holding schools and districts harmless until all components of the system are functioning effectively and validated. Flexibility is another key theme. Participants, especially those representing rural districts, want more local control and differentiated options based upon district needs, size, and location. They also support local decisionmaking regarding how much as well as when, where, and how students are tested. Funding flexibility to repair and improve school facilities, support teachers, upgrade technology, and meet other deferred needs was also mentioned. Finally, because participants feel overwhelmed and under-resourced, they desire a more gradual pace and seek to slow the roll-out of the new assessment system.

Ideal State Assessment System

When participants were asked if they could have what they want, they listed the following desired elements: a complete and balanced assessment package that includes diagnostic pre-tests, short-cycle assessments, interim and benchmark tests as well as a few, high-quality summative and end-of-course exams for key secondary classes. Differentiated assessment was another common theme. Rather than relying upon a single, point-in-time, one-size-fits-all exam to document student learning, this evaluation process would match tests and assessment methods to student readiness, interests, and preferred means of demonstrating learning. Participants suggested the need for adaptive tests and the ability to provide menus of vetted high-quality choices that align with their students' strengths, learning preferences, and other unique needs. Options for secondary students should align with course work and/or the pathways that students pursue (e.g., ACT, AP, ASVAB, IB, SAT, and industry certifications).

Unique Focus Group Themes

In addition to the cross-cutting themes, each role group articulated unique concerns. Students, for example, fear that new exams will be challenging, include unfamiliar content, and be unfair or inadequate measures for some students. They worry about fatigue given the amount of time spent viewing a computer screen during test sessions, especially for elementary students. They also want simpler tests with one section per subject area, similar to the ACT. Parents dislike the pressure placed on students, want more transparency about test items and the assessment process, and prefer fewer summative tests. Parents also seek opt-out provisions based upon individual student needs and family preferences.

Teachers reported that they are familiar with the current system and expressed fears about moving to a new and different system. They also dislike the stress that high-stakes exams place upon students and teachers. They need curricular materials, seek transparency, and want a voice in issues that affect them. Principals value the READ Act because they may select tests from an approved list, and useful, diagnostic information is obtained quickly following test administration. Principals would like to see fewer summative tests and adjustments at the elementary level (e.g., less testing and age-appropriate tests). They question the feasibility of implementing the new assessment system efficiently and need better curricular materials and test preparation guides for their staffs.

Assessment coordinators fear that new tests will not be user-friendly and view social studies exams as problematic due to timing and lack of incentives for high school seniors. This group seeks adjustments at the secondary level, such as eliminating science and social studies tests for seniors and adopting end-of-course exams. They also appreciate local choices including opt-out provisions. Technology directors hope that their districts are well prepared to administer new assessments and view feasibility (i.e., capacity, complexity, scale, resources, and timeframe) as a major challenge. They need support networks to share experiences and secure assistance; and desire additional resources to procure current and compatible devices, upgrade and maintain network infrastructure, and support additional personnel.

Superintendents raise accountability issues and concerns about how assessment results will be used to judge teachers, schools, and districts. They question the quality (i.e., length, structure, format, level of application) of new online tests. Superintendents want assurance that new tests are fair and accurate measures of student learning. They fear that the new system will be more burdensome

and require additional resources. Superintendents also view communication about the assessment process and results as a major challenge and seek incentives that will motivate students to perform well on state tests.

Finally, representatives from charter schools share the concerns raised by other groups, such as the limited value and utility of current assessments. They also mentioned that the intent of statewide testing conflicts with the values of their students' families, and hope that new English language arts and mathematics tests are developmentally-appropriate rather than requiring the same output for elementary and secondary students. The logistics are a challenge for some charter schools, especially for online schools where students and families must travel to secure testing facilities. Charter school representatives posed several solutions for easing the burden, including fewer and shorter state tests, opt-out provisions for parents, and charter school exemptions from requirements above the federal minimum.

IV. DISCUSSION

Based upon the findings and experience working with other state education agencies across the country, four implementation approaches are proposed for consideration: (1) Stay the course and implement the transition plan as scheduled; (2) Stay the course with added supports and policy adjustments; (3) Purposefully delay parts of the system; and (4) Strategically eliminate specific assessments. The fourth approach was added based upon feedback from the Commissioner's Superintendent Advisory Council. The benefits and limitations of each approach are highlighted below.

Implementation Approaches

1. Stay the Course and Implement the Transition Plan as Scheduled

Arguments in Favor:

- ❖ All components of the new state assessment program are required under state and federal statute or regulation and have sufficient educational value to merit their inclusion.
- ❖ The transition implementation plan created urgency to spur reform and incentivizes desired actions.
- ❖ Regardless of any other action taken, the number of testing components will decrease with the sun-setting of the TCAP portion of the system.
- ❖ The Colorado transition implementation plan meets or exceeds the standard achieved by most other states. Delay now might feel better in the short term but result in schools being in the same state of limited readiness at the end of any extension.

Arguments Against:

- ❖ Survey results indicate that full readiness has yet to be achieved.
- ❖ The vision of the revised Colorado Academic Standards requires a range of real reforms (e.g., curricular, instructional, assessment, professional development) that may not be achievable in the time allotted.
- ❖ The increased expectations for the use of technology to support implementation exceed current capacity (software, hardware, connectivity, and staffing).
- ❖ Too much change at once can be counterproductive.
- ❖ A new look at the current assessment system can focus on whether all components are sufficiently valid for their stated purpose, coherent across purposes, and the most efficient means to achieve its important goals.

2. Stay the Course with Added Supports and Policy Adjustments

Arguments in Favor:

- ❖ CDE, districts, and other service providers might identify additional resources and other supports to enhance readiness. For example, professional development activities might be better coordinated across all sectors to increase efficiency and coherence.
- ❖ Regional solutions, particularly in rural districts, may increase the availability of products, services, and human capital to provide differentiated support for the implementation of the various assessments and allow better use of student results for planning and targeted instruction.
- ❖ Interim cut scores could be used to mitigate harm to students and educators without delaying the overall implementation plan.

Arguments Against:

- ❖ Survey and focus group respondents indicated that time, not resources, is the biggest impediment to readiness — adding additional supports will only address part of the stated concerns.
- ❖ Consensus emerged about the value of the various components of the new assessment system. Staying the course, even with the additional supports, does not increase the perceived lack of value of specific components.
- ❖ Local assessments, while not necessarily required, were valued highly by the survey and focus group participants, and thus would not be seen as favorable candidates for easing the perceived overall burden.

3. Purposefully Delay Parts of the System

Arguments in Favor:

- ❖ A delay would allow more time for proper implementation of the valued components as well as support a renewed debate regarding the coherence of all components of the system (state and local).
- ❖ Delayed implementation of selected components of the system has two benefits. First, it respects the feedback of many educators across the state who indicated concerns with the current plan. Second, it allows CDE to remake a case for the value of the system as it currently stands or revise the system based upon additional input and information.
- ❖ As decisions are made about which components to delay, policymakers would need to consider the following items:
 - ▶ assessments required under federal statute (i.e., ESEA)
 - ▶ assessments required under state statutes and regulations
 - ▶ effects of delay of various components on at-risk student populations such as English

learners and students with disabilities

- ▶ community values which may favor content areas such as social studies or value local interim assessments over their state counterparts
- ▶ impact on college and career readiness for students in the delayed content areas

Arguments Against:

- ❖ Most components of the new assessment system are both aligned to the Colorado Academic Standards and required by state or federal statute or regulation.
- ❖ Delays of these assessments would require legislative or State Board of Education action or approval from the U.S. Department of Education.
- ❖ Delays would prevent the state from holding schools, educators, and students responsible for achieving Colorado Academic Standards in a timely manner.

4. Strategically Eliminate Specific Assessments

Arguments in Favor:

- ❖ A review of current assessments could focus on whether each one is valid for its stated purpose, and the most effective measure of student learning.
- ❖ Provides time to confirm the findings regarding the value and burden of each assessment.
- ❖ Flexibility and local options based upon district needs and size respects stakeholder feedback and could reduce the impact on instructional time.

Arguments Against:

- ❖ Tests are aligned to the Colorado Academic Standards and required by state or federal statute or regulation.
- ❖ Elimination would require legislative or State Board of Education action or approval from the U.S. Department of Education.

The implementation approaches are not mutually exclusive. The state could decide to stay the course with some assessments, delay others, phase-in the movement to online assessments, and/or eliminate specific assessments. In addition to exploring these implementation approaches, the state has identified (based on initial findings from this study) several possible short- and long-term options for feedback from districts. These options were shared and initial feedback gathered during the May 1 meeting of the Commissioner's Superintendent Advisory Council. The options, outlined below, will be shared and discussed further in phase two of the study.

Short-Term Options:

- ❖ Phase-in online assessments by allowing all math assessments to be administered by paper/pencil, thereby reducing by nearly half the amount of online testing in the first year, and by allowing those districts that have limited capacity to make a gradual transition to online assessments
- ❖ Reduce the number of high school assessments to the federal minimum:

- ▶ Make the 9th and 10th grade PARCC assessments optional for districts
- ▶ Use the culminating PARCC assessments (most students would take them in grade 11)
- ❖ Shorten the length of science assessments and only report overall scores
- ❖ Shorten the length of social studies assessments and only report overall scores
- ❖ Use a sampling approach to social studies that does not require every school to be tested every year
- ❖ Propose adjustments to the PARCC assessments:
 - ▶ shorten the number and length of sessions
 - ▶ collapse the performance-based assessments and end-of-year windows

Long-Term Options:

- ❖ Consider making school readiness assessments optional due to the perceived high burden and low value, with the noted exception of urban districts
- ❖ Initiate discussions with stakeholders to design a next generation assessment system that provides timely, relevant, and useful information that supports learning
- ❖ Explore the possibility of moving toward more authentic and instructionally-embedded assessments that focus group and survey participants valued

Study findings, implementation approaches, and potential options are being discussed among members of the Commissioner's Superintendent Advisory Council. The council's feedback informed the approaches and helped clarify options for minimizing the assessment burden on districts. Several council members expressed an interest in limiting statewide summative assessments to the federal minimum and making optional any assessments above that minimum. (See Appendix D)

The second phase of the study began in May following online administration of the new science and social studies assessments and field testing of the new online English language arts and mathematics assessments. Objectives for this second phase include the following: (1) understand whether and how the challenges and opportunities may have changed, (2) gather lessons from the state's first online administration, and (3) solicit feedback on options for facilitating a smooth transition to the state's new assessment system. The Colorado Department of Education will use the findings, input, and additional feedback to address unintended consequences of the new assessment system and inform potential adjustments to administrative policies and procedures.

Future Directions

Some of the anxiety and frustration expressed by study participants derives from how the state intends to use CMAS results for accountability purposes. Even though the study was designed to examine issues associated with implementation of the new state assessment system, some of the accountability concerns expressed by focus group participants are noted here. Besides the assessment issues identified, Colorado educators and other stakeholders raised concerns about whether the results of new tests will be sufficiently valid and reliable to support decisions about school and educator effectiveness. Focus group participants question whether new assessments are well aligned with curricula and instruction and accurately reflect student learning. And while they recognize and value the concept of growth in addition to status, some disagree with the methods used

to determine performance levels and calculate growth. Several participants also mentioned that sample sizes are too small to reflect real growth or differences among student groups, especially in small and rural districts.

Study participants stated that the current accountability system relies upon results from the same students for multiple indicators, and is punitive in nature including more sticks than carrots. They expressed the desire for a learner-centered assessment system founded upon high-quality instruction with results from key tests used as one of many measures within a balanced accountability system. Perhaps a strategy during this transition period could be facilitating an open, transparent discussion among stakeholders so that they clearly understand how much weight these new online assessments should bear in the accountability system. Such dialogues are important because the new online exams are high stakes. Results will be used to evaluate teacher performance, label schools, and judge programs.

Furthermore, the state could continue to consider the elements of an ideal statewide assessment system as well as some of the other solutions noted by focus group participants. For example, the state could simplify and adjust the timing of other major reform initiatives, specifically educator effectiveness, to ensure the stability and validity of the new assessment system prior to including results as accountability indicators. The state could initiate a conversation about sampling and how that might significantly reduce test burden while still providing state-level accountability data and serving as a proxy for growth. Finally, study participants proposed several creative solutions and incentives that are worthy of further exploration.

V. APPENDICES

A. Colorado Assessment Implementation Survey

Preparing for a new statewide assessment system presents a range of benefits and challenges for administrators, teachers, and students. To better understand these benefits and challenges, the Colorado Department of Education (CDE) is working with WestEd — a nonprofit, public research and development agency — to conduct focus groups and surveys of stakeholders as the new system rolls out. WestEd will synthesize the findings and provide a range of options to inform CDE’s planning, support, and next steps. All responses will be kept confidential.

This survey is being sent to every district assessment coordinator (DAC) in the state. DACs are encouraged to consult with other leaders in the district when completing the survey to ensure that the responses reflect those of the district as a whole. Each district should submit one completed survey.

This brief survey has three sections. Section I asks for demographic information to assist in the analysis of results. Section II asks for feedback on general readiness issues related to the new assessment system. Section III allows you to provide information on the value of state and other assessments. The survey should take 15 minutes to complete.

Questions regarding the survey should be directed to: Sheila Arredondo, sarredo@wested.org.

Section I. Demographic Information

1. **In which region of the state is your district located?**

<input type="radio"/> Pikes Peak	<input type="radio"/> Southeast
<input type="radio"/> North Central	<input type="radio"/> Southwest
<input type="radio"/> Northeast	<input type="radio"/> West Central
<input type="radio"/> Northwest	
2. **In what area is your district located?**
 - ☐ Urban (i.e., located within a major city)
 - ☐ Suburban (i.e., located within a residential region surrounding a major city)
 - ☐ Rural (i.e., located within a small town or agricultural area)
3. **How many students does your district serve?**

<input type="radio"/> 1-250	<input type="radio"/> 3,001-10,000
<input type="radio"/> 251-1,000	<input type="radio"/> 10,001-20,000
<input type="radio"/> 1,001-3,000	<input type="radio"/> More than 20,000

Colorado Assessment Implementation Study

4. Has or will your district participate in any of the following pilot activities:

- CMAS Science and/or Social Studies field tests
- PARCC ELA and/or mathematics field tests
- Educator Effectiveness pilot

Section II. Factors Affecting Readiness to Administer State Assessments

5. Please indicate the extent to which the following factors influence your district's readiness to administer required state assessments.

	Low	Somewhat Low	Moderate	Somewhat High	High	N/A
Instructional design and preparation (e.g., curricular materials and resources, instructional strategies)						
Network infrastructure (e.g., Internet access, wired and wireless connectivity)						
Devices (i.e., hardware and software at schools, device-to-student ratios)						
Management (e.g., planning, organizing, scheduling, logistics, administering, controlling)						
IT staff and personnel trained to support the process						
Procuring necessary funding and resources for modern learning environments						
Training/professional development						
Other (please specify):						

6. Please rate your district's overall readiness to administer all required state assessments.

Low	Somewhat Low	Moderate	Somewhat High	High

7. What assistance would most increase your district's level of readiness to administer the new online assessments? Please select up to three responses.

- Administration, scheduling, and logistics support
- Alignment of curricula, instruction, and assessments to Colorado Academic Standards

- Communication strategies
- Efficient methods to administer assessments while minimizing disruption to instructional time
- Procuring devices that meet instructional needs and assessment consortia requirements
- Technology specifications for the assessments
- Technology support personnel
- Training/professional development to prepare staff to administer assessments
- Training/professional development for staff to understand and use assessment results
- Other (please specify): _____

Section III. Value of State and Local Assessments

8. Please rate the value/benefit and burden (i.e., time and cost) of administering each assessment listed below.

Please use a scale where 1 is low and 5 is high.

		Value in Informing Student Progress	Value to School or District Improvement	Burden	Not Applicable
1	TCAP Reading and Writing				
2	TCAP Mathematics				
3	CMAS Science				
4	CMAS Social Studies				
5	ACCESS for ELLs				
6	Colorado ACT				
7	Early Literacy Assessments (READ Act)				
8	School Readiness Assessment (if implemented)				
9	Other district administered interim assessments (e.g., MAP, Scantron Achievement Series, STAR)				

		Value in Informing Student Progress	Value to School or District Improvement	Burden	Not Applicable
10	Other district administered postsecondary readiness assessments (e.g., EXPLORE, PLAN, PSAT, Accuplacer)				
11	Other program specific assessments (e.g., AP and IB)				

9. What does your district value most about the current state assessment system?

Please select up to three responses.

- ☐ Sets clear expectations for students
- ☐ Measures and monitors student progress toward mastery of content standards
- ☐ Documents student mastery of content standards
- ☐ Supports student placement decisions
- ☐ Ensures accountability for student performance and progress
- ☐ Provides feedback to students and families
- ☐ Provides feedback to educators
- ☐ Provides feedback to community members
- ☐ Provides common basis for state's accountability system
- ☐ Provides comparisons across schools and across districts
- ☐ Informs instructional practice
- ☐ Informs program improvement
- ☐ Documents program/school performance
- ☐ Identifies teacher development needs
- ☐ Identifies program design and delivery issues
- ☐ Other (please specify):

10. What does your district *value least* about the current state assessment system?

Please select up to three responses.

- ☐ Sets clear expectations for students
- ☐ Measures and monitors student progress toward mastery of content standards
- ☐ Documents student mastery of content standards
- ☐ Supports student placement decisions
- ☐ Ensures accountability for student performance and progress
- ☐ Provides feedback to students and families
- ☐ Provides feedback to educators
- ☐ Provides feedback to community members
- ☐ Provides common basis for state's accountability system
- ☐ Provides comparisons across schools and across districts
- ☐ Informs instructional practice
- ☐ Informs program improvement
- ☐ Documents program/school performance
- ☐ Identifies teacher development needs
- ☐ Identifies program design and delivery issues
- ☐ Other (please specify):

11. What are your *top concerns* regarding the new state assessment system?

Please select up to four responses.

- ☐ Quality of assessments
- ☐ Quantity of assessments
- ☐ Relevance of assessments to students and teachers
- ☐ Technology requirements
- ☐ Student readiness to take online assessments
- ☐ Time required for administration
- ☐ Time required for students to complete the assessments
- ☐ Impact on instructional time
- ☐ Impact on use of technology for instruction
- ☐ Timeliness of results
- ☐ Type and quality of training and preparation available to educators
- ☐ Utility of results
- ☐ Security of the assessments
- ☐ Duplication with local assessment systems
- ☐ Duplication with other state assessments
- ☐ No concerns
- ☐ Other (please specify):

12. Please indicate the importance of each element listed below in a state assessment system?

	Low	Somewhat Low	Moderate	Somewhat High	High	N/A
Actionable information at the student level						
Actionable information at the program level						
Aligned local and state assessment system						
Assessing social studies annually in grades 4, 7, and 12						
Cross-school comparisons						
Cross-district comparisons						
Cross-state comparisons						
Indicators of school readiness						
Indicators of early literacy development						
Early indicators of college and career readiness for middle and high school students						
Gradual transition from paper to online administration						
Length of assessments						
Inclusion of writing						
Information on student mastery of the Colorado Academic Standards						
Information on student growth from year to year						
Items beyond selected response (e.g., constructed responses and performance-based tasks)						
Single state assessment window						
Flexible state assessment window based upon student readiness and course progression						
Timeliness of results						
Other (please specify):						

- 13. What 3–5 essential elements would your ideal state assessment system include? In other words, if you could have exactly what you want, what would that be?**

Please limit your response to 100 words or less.

- 14. Is there any other issue not covered in this survey that you would like to share about the state assessment system? Please limit your response to 50 words or less.**

Thank you for participating in this survey.

B. Focus Group Protocol & Questions

I. Welcome and Introductions

- ❖ Welcome and thank participants
- ❖ Researchers introduce themselves
- ❖ Researchers explain that they were asked by CDE to gather data and provide an neutral, third-party summary to inform CDE's and the state's understanding of the needs and challenges facing districts as they implement the new assessment system. This study is neither a cost/benefit analysis nor a study of the Colorado Academic Standards. The focus is solely on assessment implementation issues.
- ❖ Researchers explain that while participating districts will be identified in the report, no information will be shared that could be individually identified.
- ❖ Participants introduce themselves.
- ❖ Researchers request participants to complete and submit a consent form.

II. Purpose

The purpose of the focus group is to better understand the challenges, needs, and proposed solutions for implementing a quality assessment system. CDE will use the information to inform possible changes in practices, procedures, and legislative/regulatory policy if needed.

To ensure a shared understanding of state testing, let's review what is in place now and what is changing [see one-page summary]:

- ❖ TCAP — currently includes reading, writing, math, and science
- ❖ CMAS — moving to the Colorado Measures of Academic Success in English language arts and math (PARCC-developed) as well as science and social studies
- ❖ English language proficiency assessments known as WIDA Access for ELLs
- ❖ ACT given to high school juniors
- ❖ State-mandated, locally-determined assessments for the READ Act and school readiness

Refer participants to three-page "Fact Sheet" for additional information about the changes.

We will be asking you to be specific about which assessments in this system you are referring to as well as the specific grade levels. The primary focus will be to understand the impact of the state's movement to new online assessments in English language arts, math, science, and social studies. We also want to understand how this transition is impacted by other state-mandated assessments and general reform initiatives underway.

This is a multi-phase study. We plan to conduct similar focus groups and/or interviews after the administration of the online science and social studies assessments this spring. This will help us to determine what has changed after districts administer online assessments and what new information/suggestions they might have.

Colorado Assessment Implementation Study

III. Questions

Researchers ask questions, facilitate the conversation, encourage equal air time, and take notes.

Current State Assessment System

1. What do you value most about the current state assessment system and want to retain as the state moves to new online assessments in English language arts, math, science, and social studies?
2. What do you value least and want to change about the current state assessment system?

New State Assessment System

3. What are your greatest hopes as districts implement new online assessments?
4. What are your greatest fears as districts move forward?
5. What are the top challenges that schools and districts face in implementing the state's new assessment system?
6. Moving to specifics, let's discuss some areas that may or may not have been mentioned:
 - a. Administration (e.g., scheduling, staffing, space, and logistics)
 - b. Increased expectations for students
 - c. Impact on specific populations and at-risk students
 - d. Impact on specific grade levels and subject areas
 - e. Impact on local assessment systems and uses
 - f. Impact on technology use
 - g. Impact on instructional time
 - h. Duplication of testing
 - i. Technology readiness of students and districts/schools

Managing the Transition

7. What assistance and support do schools and districts need to manage the transition to online assessments?
8. What other ideas/proposed solutions do you have for addressing the challenges mentioned?
9. What 3–5 essential elements would your ideal state assessment system include? In other words, if you could have exactly what you want, what would that be?
10. Is there anything else that you would like to share with us as we gather data and prepare a summary for CDE? [If time permits, probe further about the district's assessment system and efforts to align curricula, assessment, and instruction with standards.]

IV. Conclusion

Thank participants for their time. Inform them that information will be gathered from eight district focus groups, three role-alike focus groups, and a district survey. The aim is to develop a summary of the information by the end of April. Phase two of the work is scheduled to begin in May after administration of the science and social studies assessments. CDE will share the findings from phase one. District participation in phase two is appreciated.

V. Researcher Notes

The focus group researchers need to know the following information prior to the meetings:

- ❖ Districts that are part of the READ Act early literacy assessment tool program
- ❖ Assessments the district is using for the READ Act
- ❖ Districts that are part of the school readiness work
- ❖ Districts that participated in spring/fall field testing
- ❖ Districts participating in PARCC field testing

Handouts

- ❖ Assessment Implementation Study: An Overview
- ❖ Assessment Fact Sheet
- ❖ Assessment Summary Chart

C. Focus Group Summary Tables

I. Current System		II. New System		
Districts	Value Most	Value Least	Hopes	Fears
Archuleta	<ul style="list-style-type: none">▪ Accountability▪ Comparisons▪ ELA/math improvement	<ul style="list-style-type: none">▪ Disrupts learning▪ Does not inform instruction▪ Delayed results▪ Timing – not year end	<ul style="list-style-type: none">▪ Relevant, timely feedback▪ Less time consuming▪ Whole child focus	<ul style="list-style-type: none">▪ Cost▪ Stakes▪ Validity
Buena Vista	<ul style="list-style-type: none">▪ ACT▪ Common metric▪ Program improvement	<ul style="list-style-type: none">▪ Stakes▪ Administrative burden▪ Duration▪ Growth model▪ No student buy-in▪ Utility	<ul style="list-style-type: none">▪ Local assessment options▪ Timely results▪ Valid information	<ul style="list-style-type: none">▪ Burnout▪ Duration▪ Impact on instruction
Charters	<ul style="list-style-type: none">▪ Comparisons▪ Growth▪ Reports▪ READ Act	<ul style="list-style-type: none">▪ Quality▪ Relevance▪ Burden▪ Delayed results▪ Validity	<ul style="list-style-type: none">▪ Age appropriate▪ Adequate resources▪ Better information▪ Less time – fewer, shorter tests; immediate feedback	<ul style="list-style-type: none">▪ Impact on instruction▪ Data mining, security▪ Burden▪ Validity▪ Impact on unique groups
Cherry Creek	<ul style="list-style-type: none">▪ ACT▪ Common metric▪ Reports	<ul style="list-style-type: none">▪ READ Act – too many choices	<ul style="list-style-type: none">▪ Relevance▪ Timely results▪ Informs instruction▪ Validity	<ul style="list-style-type: none">▪ Credibility▪ Equity▪ Network infrastructure▪ Student motivation▪ Teacher evaluation
Delta	<ul style="list-style-type: none">▪ ACT▪ Common metric▪ Improvement▪ Growth	<ul style="list-style-type: none">▪ Relevance▪ Stakes▪ Student buy-in and motivation▪ Validity	<ul style="list-style-type: none">▪ Growth▪ Relevant and meaningful▪ Timely results▪ Utility	<ul style="list-style-type: none">▪ Inadequate accommodations▪ Devices▪ Impact on instruction▪ Validity

Districts	I. Current System		II. New System	
	Value Most	Value Least	Hopes	Fears
La Veta	<ul style="list-style-type: none"> Growth Status Snap shot 	<ul style="list-style-type: none"> Comparisons Delayed results Utility Validity 	<ul style="list-style-type: none"> Adaptive Aligned with content delivery Timely, useful results User-friendly, interactive, relevant 	<ul style="list-style-type: none"> Configurations and requirements Quality Stress Students' tech skills
Platte Valley	<ul style="list-style-type: none"> Data and graphs Growth Reports 	<ul style="list-style-type: none"> Disrupts instruction Cumbersome Organization Stakes 	<ul style="list-style-type: none"> Ready, prepared Student motivation Less burden Rewards 	<ul style="list-style-type: none"> Digital divide, equity Impact on unique groups Miscommunication Capacity Validity
Strasburg	<ul style="list-style-type: none"> Common metric Familiar Growth READ Act choices Reports 	<ul style="list-style-type: none"> Regulations Delayed results Student motivation Timing – not year end Impact on unique groups 	<ul style="list-style-type: none"> Growth Pilots inform policy Quality Validity – actual performance 	<ul style="list-style-type: none"> Curricula narrows Difficulty Impact on instruction Quantity – more areas Stress Student motivation Teacher evaluation

Colorado Assessment Implementation Study

I. Current System		II. New System		
Districts	Value Most	Value Least	Hopes	Fears
Woodland Park	▪ ACT	▪ CoAlt burden	▪ Engagement	▪ More accountability
	▪ Reports	▪ Frequency	▪ Interaction	▪ Increased burden
	▪ Comparisons	▪ Stakes	▪ Options	▪ Students' tech skills
	▪ READ Act	▪ Stress	▪ Timely results	▪ System failure
	▪ Utility	▪ Timing		
		▪ Utility		
		▪ Validity		
Combined Results	▪ ACT	▪ Burden	▪ Growth	▪ Equity
	▪ Common metric	▪ Delayed results	▪ Less time	▪ Impact on instruction
	▪ Comparisons	▪ Relevance	▪ Options	▪ Impact on unique groups
	▪ Growth	▪ Stakes	▪ Timely results	▪ Stress
	▪ Improvement	▪ Student motivation	▪ User-friendly, interactive, relevant	▪ Student motivation
	▪ READ Act	▪ Timing	▪ Validity	▪ Students' tech skills
	▪ Reports	▪ Utility		▪ Teacher evaluation
		▪ Validity		▪ Validity

III. Transition					IV. Future	
Districts	Challenges	Needs	Solutions	Ideal System		
Archuleta	<ul style="list-style-type: none">▪ Devices▪ Students' tech skills▪ Accommodations	<ul style="list-style-type: none">▪ Flexibility▪ Resources▪ Transparency	<ul style="list-style-type: none">▪ Facilities▪ Flexibility▪ Resources	<ul style="list-style-type: none">▪ 21st century skills▪ Balanced accountability▪ Sampling▪ Whole child focus		
Buena Vista	<ul style="list-style-type: none">▪ Bandwidth▪ Facilities▪ Impact on instruction	<ul style="list-style-type: none">▪ Professional development▪ Resources▪ Item bank	<ul style="list-style-type: none">▪ Short assessments▪ End-of-course▪ Flexibility for districts	<ul style="list-style-type: none">▪ Balanced accountability▪ Differentiated assessment▪ Multiple indicators▪ Local control		
Charters	<ul style="list-style-type: none">▪ Cost▪ Logistics▪ Staffing▪ Students' tech skills	<ul style="list-style-type: none">▪ Support▪ Teacher training	<ul style="list-style-type: none">▪ Involvement▪ Fewer tests▪ Resources▪ Sampling▪ Waivers	<ul style="list-style-type: none">▪ Flexibility, choice▪ Differentiated, options▪ Formative and interim▪ Transparency▪ Trust		
Cherry Creek	<ul style="list-style-type: none">▪ Impact on instruction▪ Capacity▪ Communication▪ Cost▪ Logistics▪ Stress▪ Timing/window	<ul style="list-style-type: none">▪ Communication▪ Flexibility▪ Autonomy	<ul style="list-style-type: none">▪ Shorter, fewer, single subject sessions▪ Rollout, thoughtful phasing▪ Waivers	<ul style="list-style-type: none">▪ Adaptive▪ Comprehensive system▪ Formative emphasis, MAP model▪ Relevant, timely feedback▪ Flexibility		

Districts	III. Transition		IV. Future	
	Challenges	Needs	Solutions	Ideal System
Delta	<ul style="list-style-type: none"> ▪ Devices ▪ Management ▪ Social studies ▪ Students' tech skills 	<ul style="list-style-type: none"> ▪ Funding ▪ Resources ▪ Professional development ▪ Transparency 	<ul style="list-style-type: none"> ▪ Lower stakes ▪ Hold harmless ▪ End-of-course ▪ Fewer tests, simplify ▪ Drop secondary CMAS 	<ul style="list-style-type: none"> ▪ Competency-based ▪ Differentiated, adaptive ▪ Formative and interim ▪ Professional development ▪ Support teachers ▪ Transparency
La Veta	<ul style="list-style-type: none"> ▪ Resources ▪ Impact on instruction ▪ Duration ▪ Capacity 	<ul style="list-style-type: none"> ▪ Bandwidth for rural districts ▪ Curricular and test preparation materials ▪ Funding and resources 	<ul style="list-style-type: none"> ▪ Flexibility, rural options ▪ Individualize, adjust timing ▪ Rollout, slow the pace, gradual staging ▪ Don't test seniors 	<ul style="list-style-type: none"> ▪ Differentiate for districts, rural options ▪ Individualized learning ▪ Focus on high-quality instruction ▪ One online system ▪ Immediate and useful feedback for educators ▪ Sampling ▪ Trust teachers
Platte Valley	<ul style="list-style-type: none"> ▪ Impact on instruction ▪ Student motivation ▪ Capacity ▪ Feasibility 	<ul style="list-style-type: none"> ▪ Materials ▪ Practice sessions ▪ Professional development ▪ Resources 	<ul style="list-style-type: none"> ▪ Comprehensive system ▪ End-of-course ▪ Incentives ▪ Timely results ▪ Don't test seniors 	<ul style="list-style-type: none"> ▪ Valid accountability framework ▪ Efficiency ▪ Flexibility ▪ Differentiation ▪ Informed decisionmaking ▪ Interim assessments

Districts	III. Transition		IV. Future	
	Challenges	Needs	Solutions	Ideal System
Strasburg	<ul style="list-style-type: none"> Cost Devices Facilities Impact on instruction Shift in purpose Seniors 	<ul style="list-style-type: none"> Curricular materials Resources Devices Network infrastructure Bandwidth 	<ul style="list-style-type: none"> Hold harmless Student consequences End-of-course Shorter, fewer tests and sections Immediate feedback Rollout, one thing at a time Timing/flexible window 	<ul style="list-style-type: none"> Appropriate accommodations Fewer standards Fewer, high quality tests Flexibility, options Individualize
Woodland Park	<ul style="list-style-type: none"> Funding for priorities Seniors Session creep Capacity Student motivation 	<ul style="list-style-type: none"> Best practices Curricular materials Exemplars Professional development 	<ul style="list-style-type: none"> Consistent, clear communication Interim/benchmark tests Flexible funding Criterion-based growth Local control Balanced stakes 	<ul style="list-style-type: none"> Shared accountability Elementary limits Invest in high-quality instruction Student-centered Incentives
Combined Results	<ul style="list-style-type: none"> Capacity Cost Devices Impact on instruction Students' tech skills 	<ul style="list-style-type: none"> Curricular materials Flexibility Funding Professional development Transparency 	<ul style="list-style-type: none"> Don't test seniors End-of-course Flexibility Rollout, thoughtful phasing Shorter, fewer tests 	<ul style="list-style-type: none"> Balanced/shared accountability Differentiation Flexibility Formative and interim

Groups	I. Current System		II. New System	
	Value Most	Value Least	Hopes	Fears
Students	<ul style="list-style-type: none"> ■ ACT (4) ■ Growth/progress (2) 	<ul style="list-style-type: none"> ■ Student motivation (4) ■ Relevance (3) ■ Validity (3) 	<ul style="list-style-type: none"> ■ Online (4) – user-friendly, interactive, engaging ■ Validity (2) – measures what students know and can do 	<ul style="list-style-type: none"> ■ Challenging content (5) ■ Students’ tech skills (4) – readiness
Parents	<ul style="list-style-type: none"> ■ ACT (4) ■ Comparisons (4) ■ Accountability (3) ■ Results (3) – longitudinal data, quality indicators ■ Utility (3) – feedback to evaluate student progress 	<ul style="list-style-type: none"> ■ Stress (5) ■ Stakes (4) – too much emphasis on one test ■ Student motivation (3) – no accountability or buy-in (secondary) 	<ul style="list-style-type: none"> ■ Accommodations (2) – appropriate, paper/pencil option 	<ul style="list-style-type: none"> ■ Cost (4) – devices, infrastructure, upgrades ■ Duration (4) ■ Impact on instruction (4)

Groups	I. Current System		II. New System	
	Value Most	Value Least	Hopes	Fears
Teachers	<ul style="list-style-type: none"> ▪ Reports (4) – data, graphs ▪ Familiarity (3) – current system and growth model ▪ Growth (3) ▪ Improvement (3) ▪ Common metric (2) 	<ul style="list-style-type: none"> ▪ Results (5) – delayed, insufficient feedback ▪ Utility (5) – too general, does not inform ▪ Burden (4) – preparation, scheduling ▪ Impact on instruction (4) – disrupts learning time, narrows curricula ▪ Stakes (4) – too high ▪ Timing (4) – too early ▪ Student motivation (3) – no buy-in or ownership ▪ Transparency (3) ▪ Validity (3) – one shot, N size, growth 	<ul style="list-style-type: none"> ▪ Results (7) – relevant, timely, user-friendly ▪ Validity (4) – growth ▪ Efficiency (3) 	<ul style="list-style-type: none"> ▪ Impact on students (12) – fatigue, stress, unique groups ▪ Impact on teachers (9) – evaluation, recruitment, retention, punitive ▪ Impact on instruction (8) – disruptive, less time, narrows curricula ▪ Students’ tech skills (4) – elementary keyboarding ▪ Validity (4) – content versus keyboarding, growth, one size for all
Principals	<ul style="list-style-type: none"> ▪ READ Act (5) – choice, diagnostic, useful ▪ Accountability (3) ▪ Reports (3) – user-friendly, strengths, weaknesses ▪ Growth concept (2) ▪ Improvement (2) – needs, programs 	<ul style="list-style-type: none"> ▪ Burden (7) – preparation, administration, CoAlt, staffing, cost ▪ Utility (4) – summative snapshot, does not inform instruction or placements ▪ Accommodations (4) – inadequate/unfair for special education students 	<ul style="list-style-type: none"> ▪ Age appropriate (3) ▪ Growth (3) – valid, useful ▪ Results (3) – timely, relevant ▪ Flexibility (2) ▪ Utility (2) – use earlier in the year ▪ Validity (2) – better aligned, actual performance 	<ul style="list-style-type: none"> ▪ Impact of instruction (8) – less time, disruptive, access for non-testers ▪ Digital divide (6) – device fear, access, skills, readiness ▪ Curricula (3) – too much to teach

Groups	I. Current System		II. New System	
	Value Most	Value Least	Hopes	Fears
Technology Directors	<ul style="list-style-type: none"> ■ ACT ■ Constructed responses ■ Growth 	<ul style="list-style-type: none"> ■ Results (3) – delayed, does not inform decisions ■ Stakes (2) – too high, heavily weighted 	<ul style="list-style-type: none"> ■ Prepared (4) – all goes well ■ Results (2) – timely ■ Appropriate accommodations ■ Choice – writing prompts ■ Efficient – less time ■ Engaging 	<ul style="list-style-type: none"> ■ Impact on instruction (3) – access to labs ■ Burden (2) – more time ■ Quality/authenticity ■ Staff capacity
Assessment Coordinators	<ul style="list-style-type: none"> ■ Utility (4) – gaps, improvement, performance indicator, interventions ■ ACT (3) – relevant ■ Common metric (2) ■ Growth (2) 	<ul style="list-style-type: none"> ■ Utility (4) – does not inform instruction, not valued by postsecondary ■ Validity (4) – N size, inappropriate measure for severe-needs students, does not reflect instruction ■ Burden (2) – administrative ■ CoAlt (2) – burden, results not useful ■ Delayed results (2) 	<ul style="list-style-type: none"> ■ High quality (4) – aligned with standards, relevant, user-friendly, valid ■ Duration – shorter tests ■ Lower burden – reduces time teachers spend preparing materials 	<ul style="list-style-type: none"> ■ Not user-friendly (5) ■ Duration (3) – too long ■ Accommodations (2) – insufficient ■ Devices (2) – insufficient numbers, access for others ■ Stakes (2) – too high

Groups	I. Current System		II. New System	
	Value Most	Value Least	Hopes	Fears
Superintendents	<ul style="list-style-type: none"> Results (6) – disaggregated feedback, snap shot, monitors mastery and progress, trends Common metric (5) – comparisons Growth (5) Accountability (4) – disaggregated data, progress Improvement (3) – math, ELA, gaps ACT (2) 	<ul style="list-style-type: none"> Validity (6) – N size, limited performance range, inappropriate for some students, proficiency cut points Quality (5) – organization, format, redundancy, not adaptive, no application Results (5) – delayed, complicated, don't inform instruction Stakes (4) – too much emphasis on one test, tied to teacher evaluation 	<ul style="list-style-type: none"> High quality (7) – adaptive, engaging, flexible, interactive, relevant, better format Results (5) – timely, relevant, useful Comprehensive system (3) – formative, summative, end-of-course Aligned with content delivery/courses (2) Validity (2) – reflects student learning, measures the right thing 	<ul style="list-style-type: none"> Burdensome (6) – more time, increased costs Accountability (5) – more for schools and teachers; none for students Validity (4) – growth, teacher evaluation, keyboarding versus content knowledge Student motivation (3) – secondary students Management (2) – planning, scheduling Students' tech skills (2) – elementary readiness Timing (2) – not year end

Groups	I. Current System		II. New System	
	Value Most	Value Least	Hopes	Fears
Combined Results	<ul style="list-style-type: none"> ACT (13) – meaningful, relevant, timely results TCAP report elements – charts, graph, cut points, disaggregated data, performance levels, trajectories, trends (10) Focus on growth (9) Common metric and Comparisons (8) Informs improvement (7) 	<ul style="list-style-type: none"> High stakes (17) – too much emphasis on a single data point Utility (16) – too general, does not guide instruction Validity (14) – cut scores, growth, knowledge, N size, not aligned with curricula Delayed results (10) Student motivation (10) – no buy-in, consequences, or accountability Timing (9) – too early; $\frac{3}{4}$ through the year Administrative burden (8) – preparing testing materials, documentation Disrupts learning and instructional time (8) Duration (7) – too many hours spent testing Relevance (7) – no value, not meaningful or informative CoAlt (6) – burdensome set up, documentation, misuse, penalties, utility Writing (5) – not reflective of teaching, poor prompt, not relevant 	<ul style="list-style-type: none"> Immediate, high-quality feedback (15) Interactive, engaging, user-friendly (12) Utility (8) – informs instruction and decisions Validity (5) – accurately reflects student learning Aligned (4) – with content delivery, courses, curricula Relevant, meaningful (4) Duration (3) – less time consuming Growth (3) – valid, useful 	<ul style="list-style-type: none"> Students' tech skills (14) – elementary readiness Duration and frequency (13) – too long, more time, extended window Impact on instruction (10) – schedule disruptions, access for non-testers Misunderstanding of results (10) – explaining low scores Student motivation (7) Validity (7) – testing technology skills not content knowledge, accurate growth Burnout (7) – student stress, fatigue Network infrastructure (7) – capacity, technical difficulties Devices (6) – inadequate and insufficient Equity (6) – digital divide, access

Groups	III. Transition		IV. Future	
	Challenges	Needs	Solutions	Ideal System
Students	<ul style="list-style-type: none"> Impact on instruction (5) Devices (3) – numbers and access for non-testers Screen time (3) 	<ul style="list-style-type: none"> Devices (4) – access to computers and labs for non-testers Technology skills (4) – time to practice and prepare 	<ul style="list-style-type: none"> Format (6) – one section/ subject, choices, web-based Frequency (3) – fewer summative, more interim Results (2) – relevant and immediate Timing (2) – test when ready 	<ul style="list-style-type: none"> Differentiated (3) – adaptive, tailored, individualized; test choices
Parents	<ul style="list-style-type: none"> Devices (4) – numbers, age/ size appropriate Capacity (2) – insufficient 	<ul style="list-style-type: none"> Professional Development Resource Bank – vetted items teachers can select Transparency 	<ul style="list-style-type: none"> Comprehensive system (6) – diagnostic, formative, interim, few summative Accommodations (5) – appropriate, individualized, paper/pencil option Teacher evaluation (4) – decrease weight of scores, student voice, growth 	<ul style="list-style-type: none"> Differentiated (4) – adaptive, individualized, multiple measures Flexibility (3) – opt out Transparency (3) – view examples, information to inform instruction Trust (3) – allow teachers to prepare students

III. Transition				IV. Future	
Groups	Challenges	Needs	Solutions	Ideal System	
Teachers	<ul style="list-style-type: none"> Digital divide (5) – impact on disadvantaged Capacity (4) – facilities, infrastructure, staff Devices (4) – availability, access for non-testers Resources (3) – limited, hard choices Accommodations (2) – appropriate 	<ul style="list-style-type: none"> Curricular materials (6) – scope and sequence, model units, resource banks Teacher tools (4) – format, question types, practice tests, navigation, troubleshooting Professional development time (3) Technology (3) – bandwidth, devices, student information portal 	<ul style="list-style-type: none"> Individualize (6) – accommodations, end-of-course, test when ready Duration (4) – short, quick, fewer sections Resources (4) – best practices, item bank, funding Sampling (4) Student accountability (4) – consequences, grades, promotion, placement Hold harmless (3) 	<ul style="list-style-type: none"> Data/results (9) – frequent, timely, relevant, useful Balanced accountability (7) – multiple indicators Individualized (7) – age and level appropriate, differentiated Respect (5) – empower and trust teachers Local decisions (4) Shared decisionmaking (4) 	
Principals	<ul style="list-style-type: none"> Feasibility (9) – bandwidth, scale, management, staffing Students’ tech skills (4) – elementary readiness Impact on instruction (3) – shut down for a month Resources (3) – taxing building resources Devices (2) – numbers, access 	<ul style="list-style-type: none"> Curricular materials (6) – textbooks, model units, exemplars Professional development (5) – best practices, support, coaching Test preparation (4) – users guides, practice sessions, remediation Management strategies (3) – efficient processes and practices 	<ul style="list-style-type: none"> Flexibility/options (8) – drop secondary CMAS; provide menu; test when ready Results (6) – timely, learner-centered, inform instruction Elementary rethink (5) – type, frequency, format Hold harmless (4) – reduce the stakes End-of-course exams (3) Interim assessments (3) Rollout (2) – one at a time Transparency (2) – teachers take the test 	<ul style="list-style-type: none"> Balanced accountability system (5) – multiple elements, indicators, and measures; compare similar groups Comprehensive system (5) – integrated; early diagnostic, more formative, interim, summative; college and career readiness, ACT approach 	

Groups	III. Transition		IV. Future	
	Challenges	Needs	Solutions	Ideal System
Technology Directors	<ul style="list-style-type: none"> Devices (6) Feasibility (6) – capacity, complexity, resources, scale, timeframe Impact on instruction (3) Management (3) – preparation, scheduling Network infrastructure (3) – bandwidth, maintenance, future planning Students' tech skills (3) Time (3) – many sessions, extended window Facilities (2) – space, compliant labs Student motivation (2) – engaging seniors, ownership 	<ul style="list-style-type: none"> Devices – compatible, comparable Network infrastructure – maintenance, consistent platform/labs Support networks – share experiences; know where to find help and resources 	<ul style="list-style-type: none"> Resources (5) – current devices, infrastructure, personnel Rollout (3) – slow the pace, one subject/year Facilities (2) – mobile labs Stakes (2) – balance the stakes, hold harmless 	<ul style="list-style-type: none"> Data (5) – accurate, high quality, informative Formative (3) – shorter, more frequent tests Multiple indicators (3) – growth, equitable comparisons Accountability (2) – balanced stakes Parent accountability (2) Portal (2) – one online system
Assessment Coordinators	<ul style="list-style-type: none"> Social studies (3) – timing, incentives for seniors Facilities (2) – space, capacity Resources (2) – limited, no money to buy new devices Timing (2) – not aligned with instruction 	<ul style="list-style-type: none"> Professional development (2) – standards, model units, technology Curricular materials (3) Test prep materials (2) 	<ul style="list-style-type: none"> Secondary (5) – drop CMAS, don't test seniors, junior year tests, end-of-course exams Rollout (4) – prep time, thoughtful phasing, support Comprehensive system (3) – pre-tests, short-cycle, interim, and summative Results (2) – timely Timing (2) – performance-based in March, summative in May 	<ul style="list-style-type: none"> Balanced accountability (5) – multiple indicators Flexibility/waivers (5) – local choices, opt-out, teacher evaluation Differentiated and relevant tests (4) – district and student options End-of-course exams (4) Hold harmless (3) – optional until 2018 Support (3) – coaching

Groups	III. Transition		IV. Future	
	Challenges	Needs	Solutions	Ideal System
Superintendents	<ul style="list-style-type: none"> Capacity (8) – sufficient processors, staff, infrastructure Impact on instruction (6) – disruptive, less time Communication (4) – explaining changes, lower scores Resources (4) – costly, hard choices, cuts Student motivation (3) – buy-in, relevance 	<ul style="list-style-type: none"> Curricular materials (5) – exemplars, model units Resources and support (4) – pilot supports Professional development (3) – instructional practices Communication and transparency (2) 	<ul style="list-style-type: none"> Flexibility (5) – for small and rural districts, reporting, frequency/number of tests, funding uses Student incentives (4) High quality (3) – fewer standards and tests Individualize (3) – adaptive, student-level mastery, match test to performance Devices (2) End-of-course exams (2) Local control (2) – less federal and state involvement Rollout (2) – slow down, assessments then evaluations 	<ul style="list-style-type: none"> Balanced, holistic, integrated system (5) – assess 21st century skills, interdisciplinary, project-based, formative, interim, summative assessments Incentives and rewards (5) – for students, parents, teachers, sites Secondary (5) – end-of-course exams, capstones, summative options, don't test seniors Elementary (3) – less testing Feedback (3) – immediate, useful; data pipeline Flexibility (3) – based on student needs, district size, location, and resources High-quality instruction (3) – basis of system Multiple measures (3) – less reliance on a single data point Sampling and spans (3) – NAEP approach Support (3) – tailored and tiered district supports Coaching and mentoring (2) – observe and monitor strategy implementation Whole child (2) – social, emotional, physical too

Groups	III. Transition		IV. Future	
	Challenges	Needs	Solutions	Ideal System
Combined Results	<ul style="list-style-type: none"> Impact on instruction (16) – lab access, preparation time, frequent disruptions, test frequency and duration Devices (11) – insufficient numbers, access for non-testers, age/size appropriate keyboards Student motivation (10) – buy-in, relevance, engagement Staffing (11) – sufficient proctors, teachers for non-testing students, analysts Management (10) – intensive preparation, scheduling, timing, extended window Resources (9) – limited, taxing building resources, daunting task, hard choices Students' tech skills (9) – inadequate, especially elementary Accommodations (7) – appropriate for young children, special education Feasibility (6) – complexity, scale Funding (6) – bandwidth, technology, substitutes Facilities/space (5) 	<ul style="list-style-type: none"> Professional development (10) – coaches, standards, assessment, curricula, instruction, technology Funding/resources (10) – technology, materials, supplies Resource banks (7) – scope and sequence, curricular materials, model units, sample items and exams Test preparation materials & practice sessions (6) – troubleshooting guides Transparency (4) Devices (2) – compatible, comparable Flexibility (2) – rural options 	<ul style="list-style-type: none"> Balanced, complete, integrated assessment system (16) – fall pre-tests, short-cycle assessments, mid-year interim or benchmark tests, year-end summative and/or end-of-course exams Hold harmless (14) – decrease weight, balance the stakes Flexibility (11) – more local control End-of-course exams (9) – no secondary CMAS Rollout (8) – slow the pace, stage logically Shorter, single-subject testing sessions (6) 	<ul style="list-style-type: none"> Shared/balanced accountability (14) Differentiated assessment (14) – individualized, choice, secondary options Feedback (11) – timely and relevant Trust and empower teachers (11) High-quality instruction (8) – data-driven, learner-centered Multiple indicators (8) Support (8) Data (7) – accurate and useful Interim assessments (6) Local decisionmaking (6) Professional development (6) College and career indicators and milestones (5) Criterion-based growth (5) Limited/minimal elementary testing (5) Relevant, high-quality tests (5) Sampling (5) Span versus level (5) Transparency (5)

D. Feedback from Commissioner's Superintendent Advisory Council

Notes from Small Group Discussions

May 1, 2014

1. What are your reactions to the scenarios that WestEd shared?
 - ❖ Scenarios look to be accurate.
 - ❖ Overall findings generally capture our scenario.
 - ❖ Stay the course
 - ❖ Stay the course — it is critical that major adjustments are not made
 - ❖ Supports and flexibility that do not impact the intent (i.e., increase in student achievement) are appreciated when districts require assistance.
 - ❖ The study affirms our sense of human resource drain to manage the shift and to maintain the necessary components of testing.
 - ❖ Consider fourth scenario of strategic elimination of certain tests
2. Based on the Phase I findings, there are a range of short- and long-term options that could be considered to address the issues raised. Possible short-term options (those that would not require changes in federal ESEA statutes) are outlined below. Please provide your feedback on the options.

Challenge	Options	Feedback
Technology readiness	<p>Phase in online assessments:</p> <ul style="list-style-type: none"> ▪ Administer all math assessments on paper for the first year ▪ Make available paper for districts with high barriers to online assessment 	<ul style="list-style-type: none"> ▪ Let those who are ready take it online ▪ No benefit: best aspect of online is that we get data in a timely manner ▪ Concerns: difficult for students to go back to paper/pencil; if results come back sooner once the roll out is completed, then this will slow down everyone if some use paper/pencil and others are online ▪ Why only for math? ▪ A possible way to get back to a short but intensive window ▪ Paper/pencil does shorten the test window. ▪ Paper as an option for those without barriers but strong values against taking away the instructional time ▪ Paper options a good alternative for any test to prevent technology issues from becoming a challenge

Colorado Assessment Implementation Study

Challenge	Options	Feedback
Access to devices	Push for emergency funds this budget cycle dedicated to device purchases (large investment in devices)	<ul style="list-style-type: none"> Similar to buying a lot of technology in a bond issue Devices result in corresponding need for human resources (support) One time investment in technology would be helpful Could be a win-win and may not only placate but also actually assist districts in tangible ways that also would enhance instruction and overall student access to what ought to be typical resources available to every child in our schools. Might manage some of the political fray Would more funding for technology be taken from somewhere else?
Length of the PARCC assessments	<ul style="list-style-type: none"> Push PARCC to collapse the performance based assessment and end-of-year assessment windows Push PARCC to shorten the test 	<ul style="list-style-type: none"> Strongly oppose the collapse: it will become the same type of assessment again Collapse assessment Shortened is not better Shorter tests Drop from three to two sections per test If they could be shortened and still remain valid then by all means that should be pushed.
Length of science and social studies assessments	Could shorten the number and length of sessions (note: would eliminate the ability to get subscore data)	<ul style="list-style-type: none"> Better than doing nothing These are critical areas that warrant tracking of student mastery Shorten maybe if we can still target critical skills and standards Eliminate these tests; no problem if subscore data unavailable Are these necessary or just preferred?

Challenge	Options	Feedback
Too many assessments	Move to a sampling approach for social studies by assessing one-third of the schools each year	<ul style="list-style-type: none"> ▪ Hesitant to advocate change from current course: highly neglected area ▪ Interesting thought: however, we do not use the current science test to inform instruction so going that route with social studies will not give us very useful data. ▪ This would help ▪ Better than doing nothing ▪ Sampling approach or eliminate ▪ Suggests that these assessments are not necessary. At the expense of instructional and learning time, if the assessments are not essential, then I do not test.
	Decrease the high school burden by making the culminating 11 th grade PARCC exams for English language arts and math required and the 9 th and 10 th grade exams optional	<ul style="list-style-type: none"> ▪ Good idea ▪ Eliminate tests: administer reading and math only in grades 3–8 ▪ Good start ▪ Agree: perhaps investigate a federal consideration of bringing 9th grade into the high school mix ▪ This makes the most sense, especially if we have a reliable interim measure used locally such as NWEA. ▪ ACT only; augment if needed ▪ Disagree: would not dispose of any
	Make the school readiness assessment optional	<ul style="list-style-type: none"> ▪ Agree: may not need mandate; currently do this when needed ▪ Good idea ▪ Important when a question exists as to a student's readiness ▪ Not needed; easy fix; change to 212 ▪ Undecided: still waiting to see whether this area has benefit or not.
Recognition of local assessments	<ul style="list-style-type: none"> ▪ Best handled through accountability changes and request to reconsider process ▪ A support the state could provide is helping districts determine what assessments are best aligned to the new standards 	<ul style="list-style-type: none"> ▪ Local assessments and their clear tie to evidence outcomes should be considered ▪ Do people like local assessments because they are familiar or because they drive improvements in achievement? ▪ Worth considering ▪ Substantiating a local assessment's use and value could be of benefit in many ways.

Challenge	Options	Feedback
Other ideas	Delay science and social studies until all issues with PARCC tests are resolved	<ul style="list-style-type: none"> Good idea: use PARCC ELA and math, get settled, and then choose to take the other two exams So much data from PARCC to focus on that science and social studies will not get much attention for awhile
	Move to the federal minimum as soon as possible	<ul style="list-style-type: none"> The increasing number of state summative assessments is reducing the good information available from local assessments, which are being put on the shelf due to the overload.

3. What input do you have for Phase II of the study? What additional data should be gathered? What questions should be asked, of whom? Where should WestEd probe deeper?

- ❖ **Fewer Tests** — Start by asking (1) what should be eliminated first? (2) how can we meet the needs of our students with fewer mandated, one-size-fits-all assessments? (3) what would districts do if they had fewer state tests? You may be surprised where the good information is coming from.
- ❖ **Impact on Instruction** — I'd substantiate anecdotal information as far as how schools and districts adjusted and the learning environment impact the changes have caused. For example, in our school we have open computer labs and it means that students in the classrooms adjacent to the open labs are not allowed to use the restroom during testing. Testing created lunch schedule issues, etc. I think the broader recognition of the impact on learning is poignant. At the same time, we don't want to throw the baby out with the bath water. There are some good potential benefits once all the problems are worked out.
- ❖ **Local Assessments** — We are planning to drop our NWEA winter assessment because it is simply too much testing and time to learn is decreased significantly with the disruption to the learning environment and scheduling for required assessments. Eliminating state tests will not reduce the amount of information collected on students. In fact, we will keep and use assessments that give information quickly.
- ❖ **Technical Issues** — Our district experienced significant technological issues during the social studies and science assessments. Although we were 100% ready to go, felt we had everything in place, and have enough devices. These tech issues now will make me question the validity of some of the assessments. Pearson will have to really ensure that all of the bugs are worked out or there will be no confidence in the assessments.
- ❖ **Technology Staff** — We had a false sense early on that maybe it was just going to be this year that we felt the urgency and the demand for additional technology staff, but we no longer feel that way. We are facing the reality that we will need a half- to full-time person to help coordinate all year long every year.
- ❖ **Testing Windows & School Calendars** — Interestingly, as we have struggled to devise a calendar for next year, the assessment windows have caused substantial concern. We did not feel comfortable limiting our testing windows, due to this year's experience with the online testing and the various things that cropped up (such as Java updates creating new challenges). For example, our "Spring" Break seemed to only fit at the beginning of March

which is about a month earlier. In addition, we delayed our school end date because we didn't want to end school with testing potentially being the last week of school. We put monthly calendars on the wall as visuals in planning our calendar and the colored indicators for test windows easily display that the last quarter of the year is extremely disrupted. That becomes a major issue for all because the "normal" of school is significantly altered.



Colorado Assessment Implementation Study

Prepared for
The Colorado Department of Education